NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

A COMPARISON OF NEURAL NETWORK AND REGRESSION MODELS FOR NAVY RETENTION MODELING

by

Bradley Steven Russell

March, 1993

Thesis Co-Advisor:	George Thomas
Thesis Co-Advisor:	Dan Dolk

Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE

1a Report Security Classification: Unclassified	1b Restrictive Markings
2a Security Classification Authority	3 Distribution/Availability of Report Approved for public release; distribution is unlimited.
2b Declassification/Downgrading Schedule	
4 Performing Organization Report Number(s)	5 Monitoring Organization Report Number(s)
6a Name of Performing Organization Naval Postgraduate School / 6b Office Symbol (if applicable) 36	7a Name of Monitoring Organization Naval Postgraduate School
6c Address (city, state, and ZIP code) Monterey, CA 93943-5002	7b Address (city, state, and ZIP code) Monterey, CA 93943-5002
8a Name of Funding/Sponsoring Organization / 8b Office Symbol (if applicable)	9 Procurement Instrument Identification Number
Address (city, state, and ZIP code)	10 Source of Funding Numbers Program Element No / Project No / Task No / Work Unit Accession No

11 Title (include security classification) A COMPARISON OF NEURAL NETWORK AND REGRESSION MODELS FOR NAVY RETENTION MODELING

12 Personal Author(s) Bradley S. Russell

13a Type of Report	13b Time Covered	14 Date of Report (year, month, day)	15 Page Count
Master's Thesis	From To	1993, March	125

16 Supplementary Notation The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

17 Cosati Codes

Field	Group	Subgroup

18 Subject Terms (continue on reverse if necessary and identify by block number)
Artificial neural networks, neural networks, reenlistment behavior.

19 Abstract (continue on reverse if necessary and identify by block number)

This thesis evaluates a possible use of artificial neural networks for military manpower and personnel analysis. Two neural network models were constructed to predict the reenlistment behavior of a select group of individuals in the Navy, from a sample of 680 individuals. The data were extracted from the 1985 DoD Survey of Officer and Enlisted Personnel. Explanatory variables were grouped into demographic/personal, military characteristics, perceived probability of civilian employment, educational level, and satisfaction with military life and military benefits. The first neural network model was compared to a more traditional method of statistical modeling (logistic regression analysis) to determine the strengths and weaknesses of the neural network model. Both models used the same set of 17 variables and were tested using a holdout sample of 100 observations. The neural network model was found to be comparable to the logistic regression model as a predictor, but deficient as a policy analysis model.

The second neural network model was constructed using the same data set and architecture as the first neural network model, including the original 17 variables, plus an additional 11 variables that consisted of variables with and without theoretical foundation for predicting reenlistment. The two neural network models were then compared and found to be similar at predicting reenlistment. Both neural network models were considered to be deficient as tools for policy analysts.

20 Distribution/Availability of Abstract X unclassified/unlimited _ same as report _ DTIC users	21 Abstract Security Classification Unclassified	
22a Name of Responsible Individual George W. Thomas	22b Telephone (include Area Code) (408) 656-2741	22c Office Symbol As/Te

Approved for public release; distribution is unlimited.

A Comparison of Neural Network and Regression Models for Navy Retention Modeling

by

Bradley S. Russell
Lieutenant, United States Navy
B.S., Southern Oregon State College, 1984

Submitted in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
March 1993

Author: ______________________
Bradley S. Russell

Approved by: ______________________
George W. Thomas, Thesis Co-Advisor

____
Daniel R. Dolk, Thesis Co-Advisor

____
Tim Hill, Associate Advisor

____
David R. Whipple, Chairman
Department of Administrative Sciences

ABSTRACT

This thesis evaluates a possible use of artificial neural networks for military manpower and personnel analysis. Two neural network models were constructed to predict the reenlistment behavior of a select group of individuals in the Navy, from a sample of 680 individuals. The data were extracted from the 1985 DoD Survey of Officer and Enlisted Personnel. Explanatory variables were grouped into demographic/personal, military characteristics, perceived probability of civilian employment, educational level, and satisfaction with military life and military benefits. The first neural network model was compared to a more traditional method of statistical modeling (logistic regression analysis) to determine the strengths and weaknesses of the neural network model. Both models used the same set of 17 variables and were tested using a holdout sample of 100 observations. The neural network model was found to be comparable to the logistic regression model as a predictor, but deficient as a policy analysis model.

The second neural network model was constructed using the same data set and architecture as the first neural network model, including the original 17 variables, plus an additional 11 variables that consisted of variables with and without theoretical foundation for predicting reenlistment. The two neural network models were then compared and found to be similar at predicting reenlistment. Both neural network models were considered to be deficient as tools for policy analysts.

TABLE OF CONTENTS

LIST OF TABLES

I INTRODUCTION

A. BACKGROUND

Military manpower and personnel analysts are continually attempting to find accurate methods of measuring manpower and personnel relationships. In today's tight budgetary environment accuracy is even more critical. For example, inaccurately predicting reenlistments could result in paying excessive reenlistment bonuses, or in having too few personnel in specific rates or ratings. Results such as these will ultimately cost the Navy money.

Manpower and personnel planners do not have accurate measures of all important manpower and personnel relationships, but they do have tools that are useful for estimating many of these important relationships. Such forecasting is primarily accomplished by using econometric models, often based on regression analysis. Depending upon their use and the level of accuracy required, these models may be simple or complex. Useful models quantify cause and effect relationships in a dynamic environment. However, it is not enough to know, for example, that an increase in the reenlistment bonus results in increased reenlistment rates. Military manpower and personnel planners must know how much a unit increase in a reenlistment bonus multiple will increase reenlistment, or how much increased advertising in a specific geographic area will increase enlistment.

One relatively new possibility for estimating important manpower and personnel relationships is the use of artificial neural networks for data analysis. Since 1990 federal agencies have spent tens of millions of dollars on artificial neural network research. The Defense Advanced Research Projects agency has spent 33 million dollars since 1990, and plans to spend another 45 million dollars to market neural network chips, develop new algorithms and test real-world applications of artificial neural networks.

Artificial neural networks applications are being explored throughout the Federal government. For example:

- The Army is testing artificial neural networks for an automatic target recognition system on the Comanche helicopter
- The Federal Bureau of Investigation is receiving bids for a prototype artificial neural network system to classify fingerprints
- The U.S. Postal Service is exploring the use of artificial neural networks for handwriting recognition.[Ref. 1]

Currently, artificial neural networks are used in areas such as securities trading, bankruptcy prediction, credit applications rating, and portfolio management. These areas are similar to manpower and personnel analysis in that they involve examining large sets of data and determining causal relationships between variables.

NeuralWare, a leading artificial neural network program, claims that artificial neural networks:

> improve the speed and accuracy of any decision that is data intensive, time intensive, and quality dependent. Neural networks can even tell you why a decision was made and what input was important. The end result is a marked improvement over conventional methods such as regression analysis, clustering, unequal promotion techniques, or other linear analysis. [Ref. 2]

Nearly all neural network programs on the market advertise that their programs are user friendly and require little or no knowledge of statistical analysis. If manufacturer assertions are true then artificial neural networks have the potential to increase the effectiveness of military manpower and personnel planners.

On February 2nd and 3rd of 1993, the first annual conference on artificial neural networks in military manpower and personnel analysis was held at the Navy Personnel Research and Development Center, in San Diego, California. This conference focused on the theory behind the use of artificial neural networks as modeling tools, current studies comparing artificial neural networks to more traditional forms of data analysis models, and future uses of artificial neural networks in military manpower and personnel analysis.

B. THESIS OBJECTIVES

The objective of this thesis is to evaluate a possible use of artificial neural networks for military manpower and personnel analysis. Recently, artificial neural networks have been receiving increased attention for a variety of research problems. However, before using artificial neural networks in the military manpower and personnel research area, they should be intensely scrutinized to determine that they are not misleading or dangerous as tools for the military analyst. In this thesis an assessment of artificial neural networks for military manpower and personnel analysis will be made and a possible use for artificial neural networks in this area will be explored.

C. RESEARCH QUESTIONS

This thesis will attempt to answer the following questions:

- Do artificial neural network programs such as NeuralWare enhance military manpower and personnel analysis?
- What are the strengths and weaknesses of an artificial neural program for data analysis?
- How does the resulting model generated by an artificial neural network program compare with a model generated by conventional data analysis techniques?

D. ORGANIZATION OF THE STUDY

The first phase of this thesis explores artificial neural networks in general. Chapter II describes what artificial neural networks are, how they operate, and in what areas they generally have been used. Chapter III reviews the literature that is pertinent to the remainder of this thesis.

The second, and analytical phase of the thesis, makes a comparison between two artificial neural network models and a more traditional model to determine the strengths and weaknesses of artificial neural networks for data analysis. Chapter IV sets out the basic methodology used in the comparison and describes the data set used to construct the models. Chapter V describes the traditional model, in this case logistic regression, used for comparison with the artificial neural network models. Chapter VI explains how the artificial neural network models were formulated to solve the chosen problem, of predicting reenlistment.

The final portion of the thesis is an assessment of the usefulness and accuracy of neural network data analysis programs for military manpower and personnel analysis. Chapter VII compares the artificial neural network models and the logistic regression model to determine the strengths and weaknesses of the artificial neural network models. Chapter VIII sets forth the conclusions about the efficacy of artificial neural networks for military manpower and personnel analysis and makes recommendations as to their further study and use.

II. NEURAL NETWORKS

A. INTRODUCTION

This chapter describes the basics of neural networks and how they function. Essentially there are two types of neural networks: biological neural networks and artificial neural networks. The human brain is an example of a biological neural network, composed of billions of neurons organized in a fashion so that it can perform complex tasks such as vision and speech recognition.[Ref. 3;p. 29] Artificial neural networks are a product of attempts to enable computers to do the types of things that the human brain does well.

Computers are high speed, serial machines designed to carry out a set of instructions, one after another, extremely rapidly. They can typically carry out millions of operations per second, which enables them to be very good at tasks such as adding long lists of large numbers. However, unlike the human brain, computers are not good at complex tasks such as pattern recognition. This is because the problem of pattern recognition is a parallel one, requiring the processing of many different items of information which all interact to form a solution.[Ref. 4;p. 3]

The early goal of neural computing was to model the human brain and to capture the underlying principles that allow it to solve complex problems. Early artificial neural networks consisted of individual electronic devices; the neurons were actual hardware in the computer. The first "neural network" was built in 1951 by Martin Minsky and Dean Edmonds. It was a large scale device that consisted of 300 tubes, motors, clutches and a gyro from a World War II bomber, all used to move 40 control knobs. The position of these knobs represented the memory of the machine.[Ref. 4;p. 47]

Today, artificial neural networks are composed of a set of computer instructions which simulates the neurons and the connections between the neurons. Information is stored as patterns, not a series of information bits as in normal computer programs. An artificial neural network does not work using a series of instructions, instead the network architecture and training method determine how the system will work. Artificial neural networks do not have separate memory for storing data; data is stored throughout the system in patterns.

1. Biological Neurons

The human brain contains approximately 10 billion (10^{10}) basic units called neurons. Each of these neurons is connected on average to about 10,000 (10^4) other neurons. Biological neurons are complicated devices that have a number of parts, sub-systems and control mechanisms. The operation of the biological neuron is a complicated and not fully understood process, but the basic details are simple. The neuron accepts inputs and adds them up in some fashion. If the neuron receives enough active inputs at once, the neuron will be stimulated and "fire;" if not the neuron will remain in an inactive state.[Ref. 4;p. 5]

A representation of the basic components of a biological neuron, the soma, the axon, synapses, and dendrites, is shown in Figure 1.

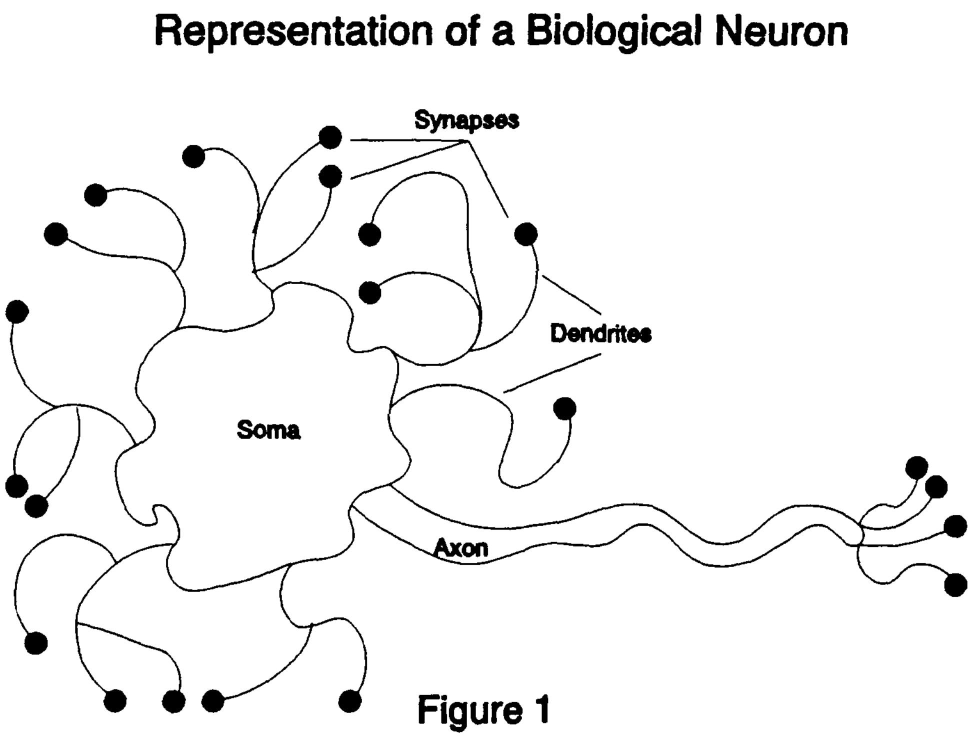


Figure 1

A brain neuron receives signals from many other neurons through **synapses**, which regulate how much of each incoming signal passes to the **dendrites**, which are the input channels to the **soma**. The soma is the body of the neuron. In the soma, incoming signals are added up and a determination made of when and how to respond to the inputs. When the neuron "fires," a pulse is sent down the **axon**, an extension of the nerve cell body. The axon is the output channel of the neuron, carrying impulses to other neurons in the brain.

2. Artificial Neurons

Artificial network neurons work in much the same way as biological neurons. A typical neuron used in artificial neural networks is shown in Figure 2. The neuron is receiving six distinct inputs from other neurons. This neuron is shown sending an output to six other neurons in the system.

Artificial Neuron Internal Representation

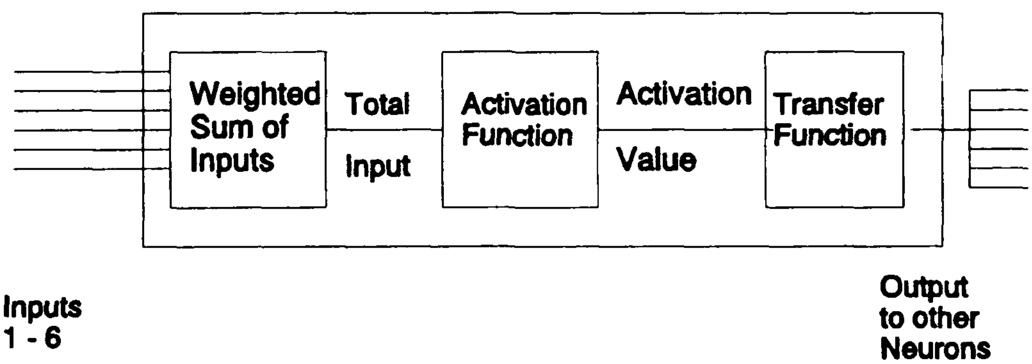


Figure 2

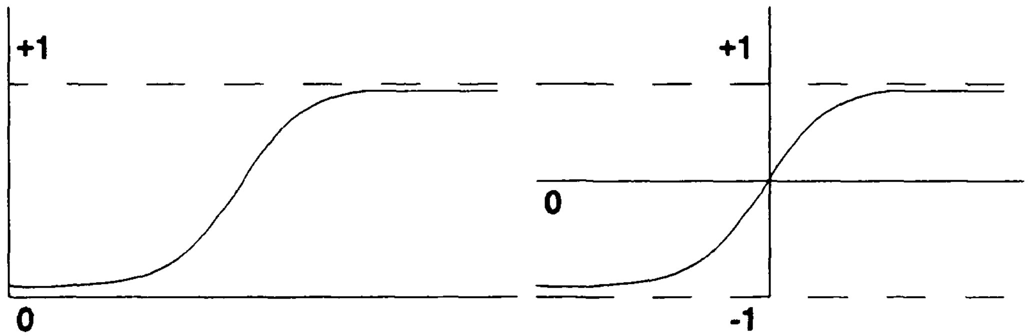
The inputs may be excitatory, tending to increase the activity of the neuron, or inhibitory, tending to decrease the neuron's activity. Once in the neuron, the inputs are weighted and combined into a single value in the box labeled **weighted sum of inputs**. Usually the inputs are simply multiplied by some weight and added together, but in some artificial neurons the calculation is more complex. Inhibitory signals can have a negative value, and thus can be added to excitatory signals but reduce the activation value. The result is the total input, which is transformed by another function know as the **activation function**.

The activation function specifies what the neuron is to do with the signals after the weights have had their effect. In the simplest models the activation function is the weighted sum of the neuron's inputs; the previous state is not taken into account. In more complicated models, the activation function also uses the previous output value of the neuron, so that the neuron can self-excite. In most artificial neural networks the activation function is deterministic, but may be stochastic in more complex networks. The activation value is then passed through the neuron **transfer function**.[Ref. 3;p. 84]

The transfer function defines how the activation value is output to the rest of the network. In some models the transfer function is a **threshold function**, or an "all or nothing" function. If the activation value is greater than some threshold amount then the neuron will output a one; conversely an activation value less than the threshold value will result in a zero output. In this model the neuron's activation must reach a certain level before the neuron adds to the total network state.

Most common artificial neural networks use a transfer function known as the **saturation function** in which more excitation above some maximum firing level has no further effect on the output of the neuron. Examples of saturation functions that are widely used in artificial neural networks today are the **sigmoid function** and the **hyperbolic tangent function (Tan H)**. These functions yield output which is a continuous, monotonic function of the input. Both the functions and their derivatives are continuous everywhere, and their values asymptotically approach a high and low value, with a smooth transition in between. The sigmoid transfer function's output (shown in Figure 3) approaches zero when its input is a large negative number, and approaches one when the input is a large positive number. The Tan H transfer function's output (shown in Figure 3) approaches negative one when its input is a large negative number, and approaches positive one when its input is a large positive number. The sigmoid transfer function is typically employed in those networks which are used for classification, while the Tan H transfer function is used in those networks involved in prediction.[Ref. 3;p. 87]

Common Transfer Functions



Sigmoid Function

Hyperbolic (Tan H) Function

Figure 3

Artificial neurons are sometimes compared to latches. A latch is a digital circuit with a feedback loop which causes it to retain or store its state. A latch can hold that piece of data indefinitely. Neurons do not hold specific on/off information, instead they keep track of how they respond to the neurons connected to them and fire based upon their input. When a neuron fires it sends out a signal. The length of time spent firing a signal is constant but the overall firing frequency is variable. Higher firing frequencies signal that the neuron is more excited.[Ref. 3;p. 19]

B. CHARACTERISTICS OF ARTIFICIAL NEURAL NETWORKS

1. Terms and Definitions

Many types of artificial neural networks exist today. It is beneficial to understand some of the terms that define and describe different types of neural networks before discussing them in detail. Various terms and simple definitions that describe behavior and abilities are presented in the remainder of this section.

Adaptability is the ability to modify a response to changing conditions in the network. Four separate processes produce this ability: Learning, training, self-organization, and generalization. **Learning** is the process by which a network modifies its connection weights in the activation function of the neuron. There are two types of learning: supervised and unsupervised. **Supervised learning** is characterized by an outside influence (either a set of training facts or an observer) telling the network whether or not its output is correct. The network's output is compared to the correct output, and the synaptic weights in the individual neurons are adjusted to make the next output closer to the desired output. In **unsupervised learning** the network does not use a set of training facts nor is it coached by an outside observer. Rather, it classifies inputs as patterns that share common features with other input patterns, with no regard to actual output.[Ref. 3;p. 88, 219, and 223]

Training is the process in which the connection weights are modified in some fashion, using the learning method. **Self-organization** is how artificial neural networks train themselves according to the learning rule. Typically all of the network's neuron weights are modified at the same time.

Generalization is the network's ability to classify patterns that have not been previously presented to the network. Networks generalize by comparing input patterns to the patterns held in the synaptic weights of the individual neurons. A pattern that the network previously has not seen is classified with other patterns that share the same distinguishing features as those on which the network has been trained.

In typical computers, if a sector of memory is lost, the program will fail. However, an artificial neural network will continue to function, but at a reduced speed and capacity. **Plasticity** is the ability of a group of neurons to adapt to different functions over time. When a portion of the network is damaged, other neurons adapt to take over functions that the damaged portions performed. **Fault tolerance** is the ability to keep processing, at a reduced speed and capacity, when a portion of the network is damaged.[Ref. 3;p. 88]

Most training data sets will typically have outliers in the data, that is, observations that are outside the "normal" range for the set of observations. **Dynamic stability** is the ability of the network to be given an extreme observation and yet remain within its functional boundaries and reach a stable state. **Convergence** is the changing state of the network as it moves towards that steady state.

2. Layers

A neural network consists of groups of neurons arranged in structural units known as layers. A layer of neurons is a group of neurons that share a functional feature. There are three possible types of neurons in a neural network, each type relating to the layer in which it lies in the network. The **input layer neurons** receive data from the outside world, from data files, keyboards or other transmitting devices. The **output layer neurons** send information back to the user in a form defined by the setup of the network. The **hidden layer neurons** are all of the neurons lying in the layer(s) between the input and output layers. Neural networks may have only one hidden layer, no hidden layers, or many hidden layers, depending on the architecture and complexity of the network and the computing capacity of the user computer. The user will not see the inputs and outputs of the hidden neurons because they connect only to other neurons.[Ref. 3;p 79]

3. Network Architecture

Artificial neural networks fall into one of two basic network architectures, **feed forward** and **feedback**. Feed-forward networks have two or more layers, each of which receives input from the preceding layer, and sends output to the succeeding one. These types of networks have no connections between neurons in the same layer. Each neuron in one layer is connected to every neuron of the succeeding layer. Thus, the network only feeds information forward in the network to the next layer of neurons. Feed-forward networks compute results very quickly because there is no delay while the neurons interact with each other and settle into a steady state.[Ref. 4;p. 7-9] An example of a feed-forward neural network is shown in Figure 4.

In a feed-forward network, results are computed by first entering values to the input neurons. The input neurons calculate their output values which are passed to the hidden layer neurons. Each hidden neuron sums the values of the input neurons, based on the weighing factor of each separate hidden neuron. The connection weights, stored in the activation function, comprise the knowledge stored in this type of artificial neural network. These connection weights correspond to the synapses in biological neural networks. When the hidden neurons are finished computing their results, they are passed to the output layer neurons. The output neurons compute their results in the same manner, based upon the weighted sum of the signals from the hidden neurons.[Ref. 3;p. 153]

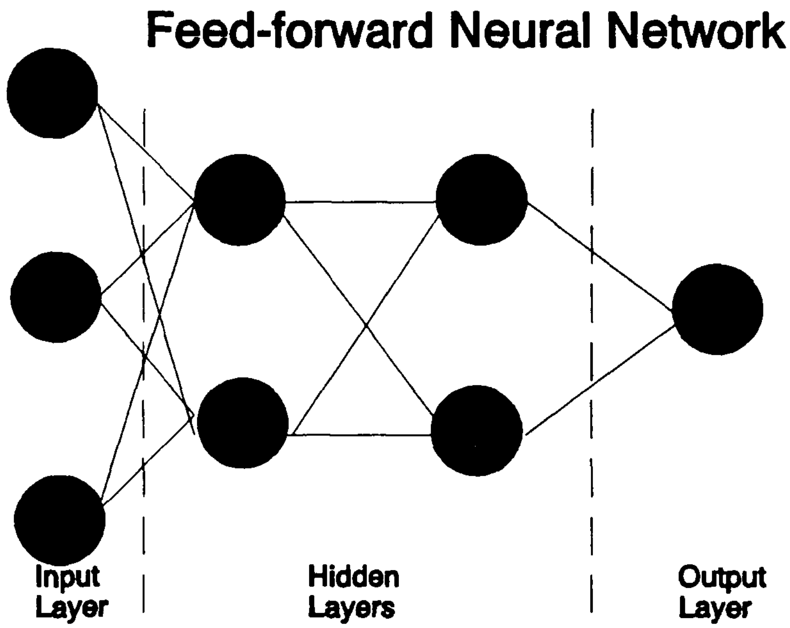


Figure 4

Feedback networks are characterized by neurons which take their inputs from any other neuron, even from themselves. Inputs are given to the network and the results are computed repeatedly until the network neurons settle into a stable state. Feedback networks are good at reconstructing facts from incomplete and error filled inputs.

4. Network Classification and Description

This section explains the various classifications of artificial neural networks shown in Figure 5, and briefly explains the theories behind the networks. Because this thesis uses the backpropagation learning algorithm as its basic artificial neural network, much of the remainder of this section is devoted to backpropagation and its predecessor, the

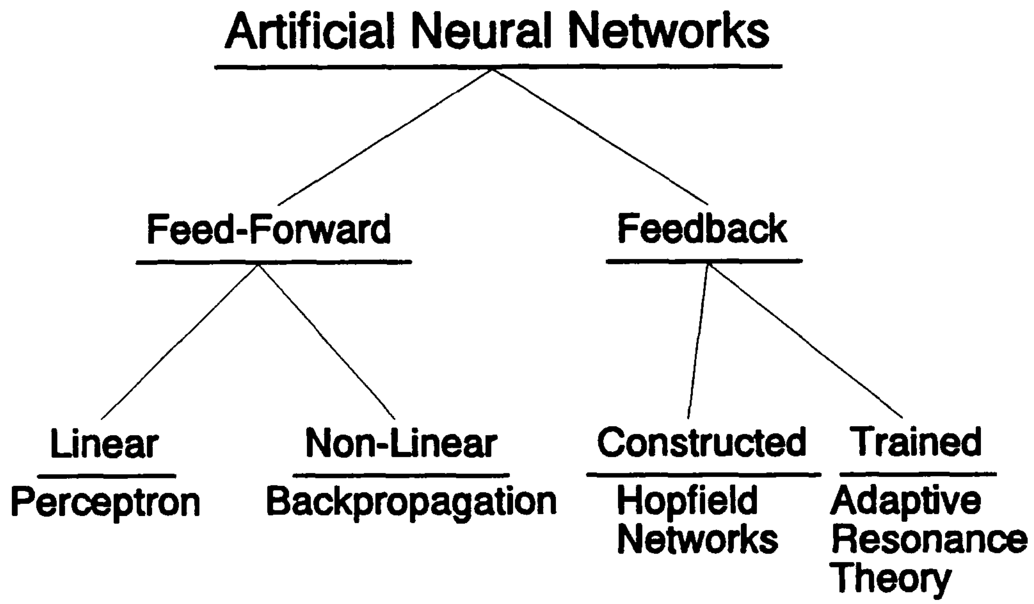


Figure 5

perceptron. A basic mathematical foundation for these types of artificial neural networks is provided. The remainder of this section provides a short description of other artificial neural networks not used in this thesis, but used in other areas today.

a. Perceptrons

The perceptron, developed in 1957 by Frank Rosenblatt of Cornell University, was the result of one of the first major research projects in the field of artificial neural networks. A simple perceptron neuron with two inputs and one output is shown in Figure 6. The term X_0 is always positive one, and the weight W_0 is referred to as the bias, and operates like the constant in a regression equation.

Simple Perceptron Neuron and Step Transfer Function

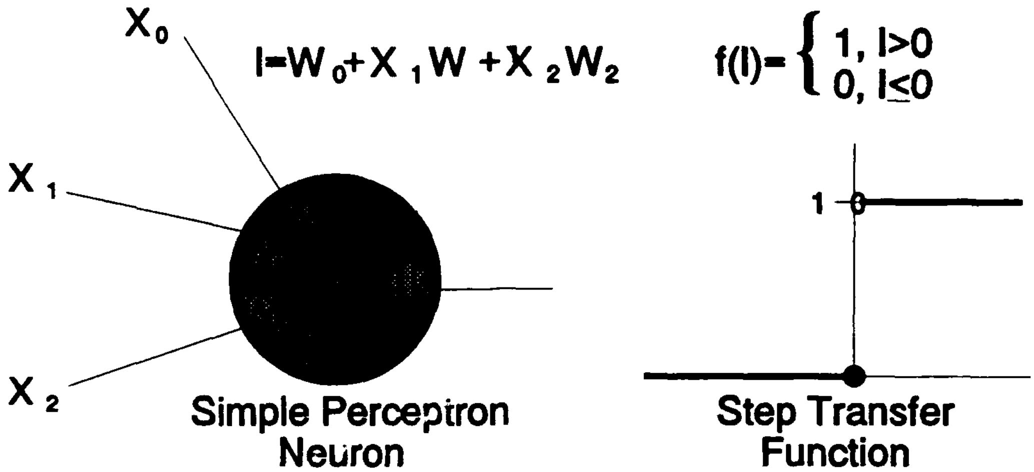


Figure 6

The perceptron network is essentially a linear separator. If we assume a simple network with two neurons in the input layer and one neuron in the output layer, the network can be used to separate the two classes of output shown in Figure 7.

When the network begins with random weights, occasionally the inputs to the network will result in a correct output. However, some of the input combinations will result in incorrect outputs. In these cases the weights need to be adjusted so that future sets of inputs will yield correct outputs. This adjustment of weights is referred to as learning. The learning algorithm for the perceptron network, as modified by Widrow and Hoff in 1960 follows:

Two Linearly Separable Classes

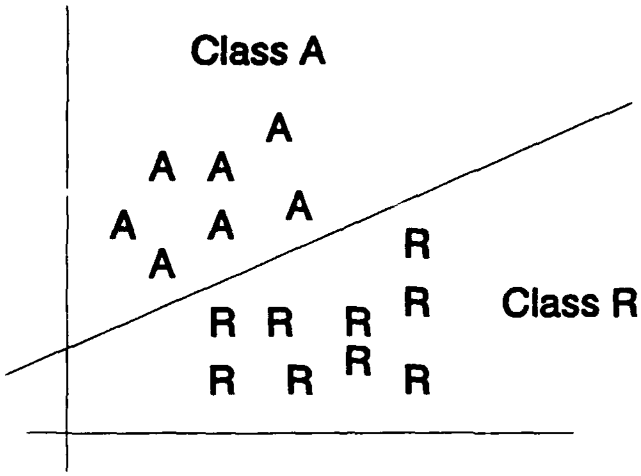


Figure 7

0. Randomly initialize the weights and the bias

1. Present an input pattern $(X_{1t}, X_{2t}, \ldots, X_{nt})$ and a desired output d_t to the network

2. Calculate the actual output of input t, y_t, from the network: $y_t=f[\sum X_{it}W_{it}$

3. Compute the error of output t, e_t: $e_t=d_t-y_t$

4. Compute the new weights for input t+1: $W_{it+1}=W_{it}+\alpha e_t X_{it}$ where α is the learning rate, $0<\alpha<1$

5. Repeat steps one through four for each new input pattern $(X_1, X_2, \ldots, X_n)$

6. Repeat steps one through five until error is less than some preset tolerance.

For the above example $d_t=1$ if the desired output is from class A, and $d_t=0$ if the desired output is from class R. If W_1 and W_2 initially are randomly set to one and the bias is set to zero, the initial line will have a slope of negative one and an intercept of zero. As the perceptron is fed input patterns and learning is accomplished through the Widrow Hoff delta rule, the line separating the two categories will gradually shift until the slope is equal to $-X_2/X_1$, and the intercept is equal to $-W_0$. This gradual shifting of the linear separator is shown in Figure 8. Line one (L1) is the beginning line, with initial weights of positive one, and line five (L5) is the hypothetical ending line that the network produces that separates class A from class R.

Two Linearly Separable Classes

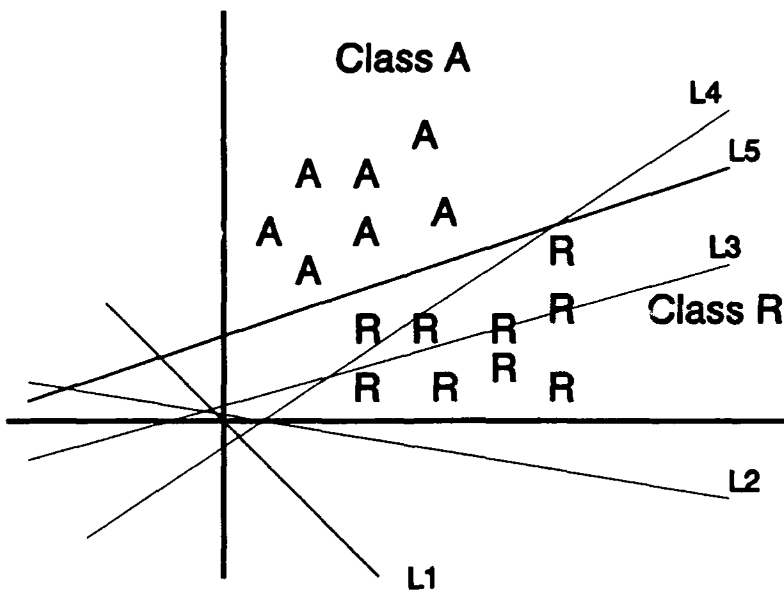


Figure 8

As previously stated, the perceptron was the result of early work in the field of artificial neural networks. As with any model, the perceptron has limitations to its capabilities. It will learn a solution if the problem is linearly separable. In many cases however, the separation between classes is much more complex. The classic simple problem that the perceptron is unable to solve is the case of the exclusive-or (XOR) problem. The XOR logic function has two inputs and one output. It produces an output only if either one or the other of the inputs is on, but does not produce an output if both inputs are off or both inputs are on. The exclusive-or problem is shown in both tabular and graphic form in Figure 9.

Exclusive-Or Problem

X_1	X_2	Y
0	0	0
0	1	1
1	0	1
1	1	0

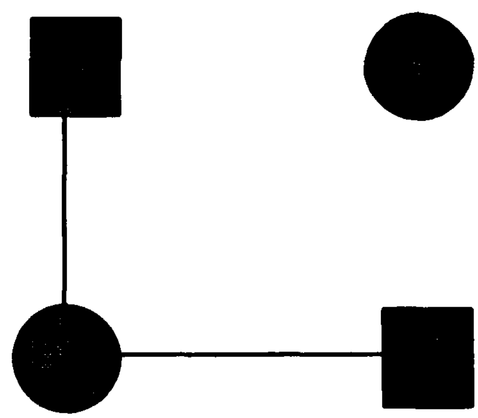
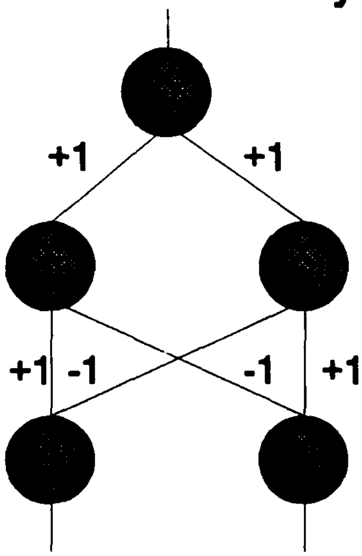


Figure 9

The logical sequel to the simple perceptron was a multi-layer network of simple perceptrons. Intuitively it can be seen that a multiple layered network with the right weights would be able to solve the XOR problem. Such a network, with the correct weights to solve the XOR problem, is shown in Figure 10.

Recoding XOR into a Lineraly Separable Problem



X_1	X_2	X_3	X_4	X_5
0	0	0	0	0
1	0	1	0	1
0	1	0	1	1
1	1	0	0	0

Figure 10

The drawback to this network is that the weights must be correctly set or "hard coded" so that the input data is mapped into a linearly separable space. If the weights are randomly set at the start, the network will be unable to learn. This is because there is a credit assignment problem inherent in a multi-layer network with neurons that have a step transfer function. The "on" or "off" state of the neurons give no indication of the scale by which the weights need to be adjusted for incorrect output. The step transfer function thus removes the information about the input that is needed if the network is to learn.[Ref. 4;p. 65]

Minsky and Papert in Perceptrons [Ref 5] pointed out the limitations and criticisms of single and multiple layer perceptron networks. They demonstrated that perceptrons could only do linearly separable problems; this was the "brick wall" that the artificial neural network field of study ran into in the 1960's. During this time however, large strides were being made in the field of artificial intelligence, solving many of the problems that perceptrons could not. Thus gradually most of the major funding shifted from the study of artificial neural networks to artificial intelligence during the following twenty years.

Relying heavily on pre-processing inputs to form nearly linearly separable sets of data, perceptron artificial neural networks have been used in various applications. These include research of speech recognition, character recognition and adaptive noise filtering. Also, in Japan a university researcher has used a perceptron artificial neural network to build robots that have learned to walk.[Ref. 6]

b. Backpropagation

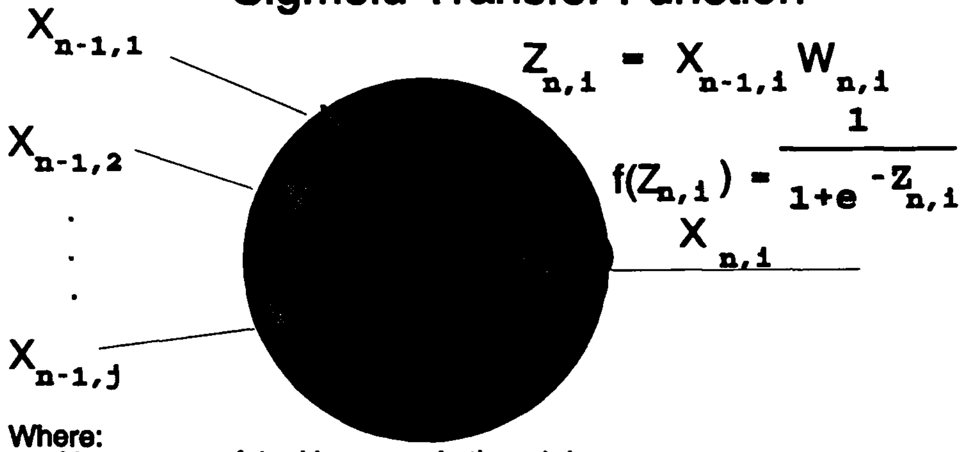
In 1986 a breakthrough in the study of artificial neural networks was put forth by Rumelhart, McClelland, and Williams in their book Parallel Distributed Processing [Ref 7]. Their breakthrough was a way to use a smooth transfer function in a multi-layer perceptron network, combined with a learning rule which "backpropagated" the error from the output layer to the input layer, thus solving the credit-assignment problem.

The term "backpropagation" refers to a type of learning algorithm for adjusting the weights in a multiple layer feed-forward network. However, the term has become synonymous with the type of network itself, and will be used in this context for the remainder of the thesis.

In backpropagation, the responsibility for output error is assumed to be the problem of all the connection weights in the network. Errors are calculated at the output layer, then using a sum of products to the previous layer, the previous artificial neurons are assigned error. The errors are then used in adjusting the incoming weights so as to produce an output closer to the correct output for the next set of learning inputs.[Ref 6]

Two of the most common transfer functions used in backpropagation are the sigmoid and the Tan H transfer functions discussed earlier in this chapter. These transfer functions have relatively simple, continuous derivatives. These derivatives are the basis for the backpropagation learning algorithm; they are used to assign error to each of the artificial neurons in the network. An artificial neuron that uses the sigmoid transfer function is shown in Figure 11.

Backpropagation Neuron Using a Sigmoid Transfer Function



Where:

$X_{n,i}$ = output of the ith neuron in the nth layer

$W_{n,i,j}$ = weight of the output of the jth neuron in the (n-1)st layer to the ith neuron in the nth layer

Figure 11

The general procedure for backpropagation follows:

0. Initialize weights, $W_{n,i,j}$, randomly

1. Present an input pattern $(X_{1t}, X_{2t}, \ldots, X_{nt})$ and a desired output d_t to the network

2. Calculate the actual output for the input pattern $(X_{1t}, X_{2t}, \ldots, X_{nt})$, y_t, from the network: $y_t=f[\sum X_{it}W_{it}$

3. Compute the total sum of squares error for the network for input t, e_t: $e_t=0.5*SUM_t(d_t-y_t)$

4. Calculate $\Delta W_{n,i,j}$ (Described in following paragraphs)

5. Feedback: Correct the weights $W_{n,i,j}(new)=W_{n,i,j}(old)+\Delta W_{n,i,j}$

6. Repeat steps one through five for all training patterns

7. Repeat steps one through six until the error is less than some pre-determined tolerance.

The basic formula for changing the weights is:

$$\Delta W_{n,i,j} = alpha * X_{n-1,i} * e_{n,i}$$

where: $X_{n-1,i}$= output from neuron i of layer n-1

$e_{n,j}$= error of neuron j in layer n

alpha = learning rate, 0<alpha<1

There are two formulas for calculating a specific neuron's error. The formula for a neuron's error in the output layer is directly proportional to the difference between the desired output and the actual output of the output neuron. It also depends on the derivative of the transfer function for the neuron in the output layer. This formula is:

$$e_{j,out} = f'(Z_{j,out}) * (d_j - y_j)$$

The formula for a neuron's error in any layer below the output is proportional to the backpropagated error. This means that the error in these nodes depends on the errors of the nodes above and the connecting weights to the above nodes. The neuron's error in any layer below the output layer also depends upon the derivative of its transfer function at its current output level. This formula is:

$$e_{j,n} = f'(Z_{j,n}) * SUM(e_{k,n+1} * W_{k,j,n+1})$$

Thus, the change in an incoming weight is proportional to the error of a neuron times the value of the input on the connection being adjusted.

One modification to the backpropagation procedure, developed to avoid local minima in the error structure is the "generalized Delta rule." This modification adds a momentum term to the change in the $W_{n,i,j}$'s This momentum term is a constant, β, multiplied by the weight vector of a neuron from the previous presentation of an input pattern, which is then added to the next change in the weights to avoid local minima in the error structure. The new formula for changing the weights by the generalized Delta rule is:

$$\Delta W_{n,i,j} = \text{alpha} * X_{n-1,i} * e_{n,i} + \beta [W_{n,i,j(old)} - W_{n,i,j(new)Prev}$$

Backpropagation is thus able to solve the XOR problem because outputs from the neurons can take on intermediate values between either zero and one (for the sigmoidal transfer function), or negative one and positive one (for the Tan H transfer function). This allows a network to slowly readjust its weights in the individual neurons, and to move down the error structure until some preset error tolerance level is reached.

The number of applications for multiple layer, backpropagating artificial neural networks is continually increasing. Some of the areas in which they have been used are sonar interpretation, machine vision, converting english text to phonemes, airline seat marketing, and forecasting in the economic and banking areas. They have applications in pattern classification, modeling complex non-linear functions, and signal processing problems. Additionally, they are beginning to see wide use in the field of robotics.[Ref. 7]

c. Hopfield Networks

Hopfield networks are fully-connected feedback networks. They consist of a number of neurons, each connected to every other neuron in the network. They are symmetrically weighted networks, each link from one neuron to another having the same weight in both directions.

The Hopfield Network

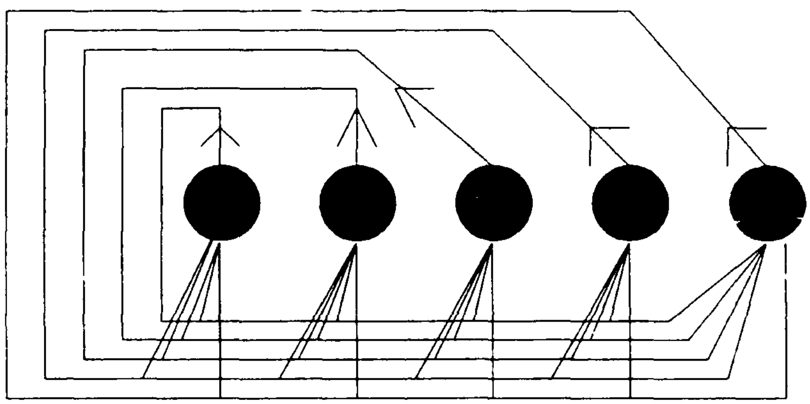


Figure 12

Figure 12 shows a fully connected Hopfield network. The major distinguishing feature of the network is that there are no obvious input and output neurons, and this architecture defines how the network will operate. Inputs to the network are applied to all of the neurons at once, consisting of a set of starting values, either positive one or negative one. The network is allowed to cycle through a succession of states until it converges on a steady state solution (if one exists!). This steady state occurs when the values of the neurons no longer change. Because each neuron is connected to all other neurons in the system, the output value of one neuron affects the value of all others. The initial, unstable state is characterized by many different values each affecting each other. As the net moves through a succession of states it is trying to reach a compromise between all the values in the network, and the final steady state represents the solution to the inputs. In this state there are as many inputs trying to turn on a neuron as there are inputs trying to turn it off, so it remains in a stable, steady state.[Ref. 4;p. 133-135]

Hopfield networks have seen limited commercial applications because of the relatively short amount of time that researchers have been working in this area. Hopfield networks have applications in the field of simulated annealing, or the process used to improve the characteristics of crystals or metals. Because of their high tolerance of partial damage to the network, Hopfield networks hold great promise in the field of space-based electronic and robotics systems, where radiation damage to computer chips is a possible occurrence.

d. Adaptive Resonance Theory

The adaptive resonance theory is a two-layered, feedback network type. The major feature of the adaptive resonance theory is the ability to switch from a plastic mode, where internal parameters of the network can be modified, to a stable mode where the internal mechanics of the network are fixed, without losing any previous learning.

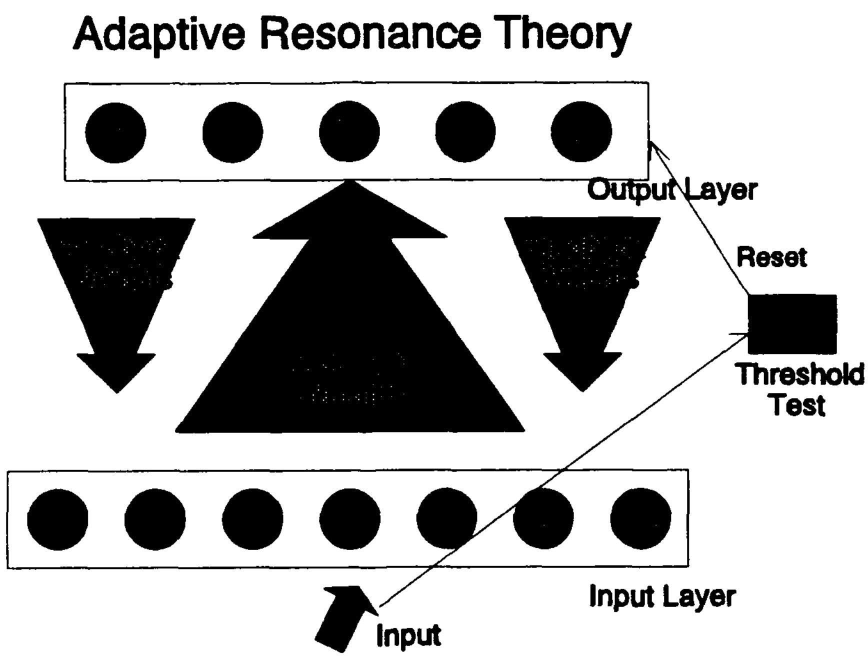


Figure 13

An adaptive resonance theory network, shown in Figure 13, has two layers whic'. are connected with extensive use of feedback. Feedback flows from the output layer to the input layer, and also between neurons in the output layer. An adaptive resonance theory network is a combination of a feed-forward network, and a feedback network, but is classified here as a feedback network because of its extensive use of feedback not found in other types of feed-forward networks.

For each layer there are logic control circuits that control the movement of the data through the layers at each stage of the operating cycle. Between the input and output layers there is a reset circuit responsible for comparing the inputs to a threshold that determines whether a new class pattern should be created for an input pattern.[Ref. 4;p. 167-169]

Adaptive resonance theory is a self-organizing network that has been able to solve the stability-plasticity dilemma, and has been applied to several pattern recognition problems in a laboratory setting. Adaptive resonance theory networks have not been used in commercial applications, probably due to the newness of the theory.

C. OPERATION OF A NEURAL NETWORK

The normal operation of a neural network is a selective response to a signal pattern. How each specific network learns is determined by type of connections between the neuron, the weight assigned to a signal, and the rules which change the input function.

An example which helps to explain the operation of a neural network is that of a network trained to predict a dependent numerical output from a set of inputs, or explanatory variables. A feed-forward, backpropagating network is used in this case. Each of the explanatory variables is assigned to an input neuron, which in turn sends signals to the next layer of neurons, the hidden layer. Each hidden neuron receives signals from all the neurons in the preceding layer. The signals are assigned connection weights and summed in the activation function of the neuron. If the activation value is greater than the threshold value, the neuron "fires" and sends a signal to the next layer. If less than the threshold value, the neuron remains in an inactive state. Once all of the inputs have been passed through the hidden layer the outputs are sent to the output layer of neurons.

The output layer of neurons, in this case only the one neuron associated with the dependent variable that is being predicted, is compared to a value known as the training value. The training value is the actual value of the dependent variable for the explanatory variables in the observation. In the back propagation learning method the predicted value is compared with the actual value of the dependent variable, and if there is a difference, an error signal is fed back throughout the network, altering the connection weights in each of the neuron's activation functions. The network iteratively moves to the next observation in the data set, until a pattern is formed and the network can successfully predict and match all of the output values to their actual values.

At this point the network is considered trained and ready for testing by the user. Testing is accomplished in much the same manner as training. A separate testing data set with new explanatory and dependent observations is input into the network. The predicted outputs are compared with the actual dependent values to determine how well the network is performing on data separate from the training data set.

The next chapter presents a review of the pertinent literature that compares the use of neural networks to more traditional methods of statistical modeling.

III. LITERATURE REVIEW

A. INTRODUCTION

The prediction of manpower and personnel behavior is a necessity in the military decision making process. Typically these predictions are made using some type of multiple regression model, with cause and effect relationships hypothesized between the independent and dependent variables. However these regression models have various problems associated with them. First and foremost is the choice of the underlying functional form of the model. If the researcher incorrectly specifies this initial formation of the model, the model will be much less likely to perform well as a predictive tool. Other problems with regression are the assumptions that must be made in order for regression to be a valid prediction technique. Normality and independence of the error term, and constancy of the error variance are assumptions which are often made (and frequently not tested) when using regression models.

Neural networks allow predictive models to be created without a priori knowledge of the functional form. Assumptions about normality, independence, and constancy are not required in the neural network model. For these reasons, neural networks should be examined to determine their efficacy as a tool for helping the military decision maker.

The use of neural network models as a tool for analyzing data sets is a relatively new field. The development of the error backpropagation learning algorithm, by Rumelhart, McCelland and Williams in 1986 [Ref. 6], opened the research area for many new applications of neural networks. However, only a small number of researchers have compared the use of neural networks to traditional data analysis techniques in the area of military manpower and personnel research. The recently held, first annual conference on neural networks in military manpower and personnel analysis at NPRDC highlighted awareness in the field that neural networks are a new modeling tool that needs to be evaluated. This thesis is an effort to provide an evaluation of neural networks as a modeling tool for the military manpower analyst.

This chapter reviews the pertinent literature comparing neural networks and traditional military manpower and personnel modeling techniques. In addition, it reviews other literature which compares neural network models with multivariate and bivariate analytical techniques in the fields of bankruptcy prediction, bond rating, and stock price predictions. These areas share many characteristics with military manpower and personnel analysis. Both manpower and personnel analysis, and economic analysis typically involve the interaction of many unrelated variables, making prediction difficult and complex. For this reason, it is worthwhile to review the results of studies comparing neural networks to traditional data analysis techniques in fields other than military data analysis.

B. COMPARISONS OF NEURAL NETWORKS AND CLASSICAL FORECASTING METHODS IN THE MILITARY

Dickieson and Wilkins [Ref. 8] compare neural networks with multiple regression in the prediction of premature attrition from the U.S. Naval Academy. Both types of models were developed using the same seven explanatory variables currently in use by the Naval Academy.[1] The dependent variable for the study, voluntary attrition, is dichotomous. The study uses the data of three recent classes from the academy, referred to in the study as classes I, II, and III. [Ref. 8;p. 67]

The regression model used for this study is based on stepwise ordinary least squares (OLS) regression, essentially the same model now used by the academy. The model is estimated using data from class I, then cross-validated using data from class III. The correlation between predicted attrition and actual attrition in this model was found to be .0561. The authors explain that the correlation coefficient

[1]These variables are SAT-verbal, SAT-quantitative, high school rank in class, recommendations from high school officials, extracurricular activity score, technical interest score, and career interest score.

is small because attrition is difficult to predict, it is a dichotomous variable, and because few people actually are prematurely discharged from the Naval Academy.[Ref. 8;p. 68]

The construction of neural networks is often described as more of an art than a science. Choices must be made as to what type of architecture to use, the number of hidden layers, and number of neurons in each hidden layer. This study uses six different neural networks to determine their impact upon whether neural networks outperform regression in predicting attrition from the Naval Academy. Table 1 shows the various characteristics of these models.

TABLE 1: NEURAL NETWORK'S CHARACTERISTICS

Network	Architecture	Inputs	Hidden Layer 1	Hidden Layer 2	Outputs
1	Backpropagation	7	14	0	1
2	Backpropagation	7	7	0	1
3	Functional Link	7	7	0	1
4	Functional Link	7	4	3	1
5	Backpropagation	7	21	0	1
6	Backpropagation	7	2	0	1

Source: Dickieson and Wilkins (1992)

In developing neural network models for this problem, two different stopping criteria are used. The six neural network models are developed using data from Class I, and then cross validated on Class II data to determine the separate stopping criteria. Criterion A is the number of iterations which produced the maximum cross-validation correlation coefficient between predicted and actual attrition. Criterion B is the midpoint of the range of iterations for which the neural network model outperformed the linear regression model for Class II data.[Ref. 8;p. 69]

After the two stopping criteria are developed, the six neural network models are cross validated on the Class III data to determine the predictive efficacy of the models. For all six networks, criteria A and B yield correlations higher than those provided by linear regression. The results of both the neural network models and the linear regression model are shown in Table 2.

TABLE 2: CLASS III CROSS-VALIDATED CORRELATION COEFFICIENTS

Network	Regression	NN-Criterion A	NN-Criterion B
1	.0561	.0846	.0806
2	.0561	.0806	.0762
3	.0561	.0854	.0858
4	.0561	.0577	.0577
5	.0561	.0860	.0759
6	.0561	.0657	.0657

Source: Dickieson and Wilkins (1992)

The results of this study show that neural network models can have a higher predictive efficacy than stepwise linear regression. However a more plausible regression model may have yielded better results. In light of the dichotomous dependent variable, a logistic form of model rather than a linear model may have yielded a higher correlation between predicted and actual attrition.

Wiggins and Engquist [Ref. 9] compared neural networks to probit regression analysis in predicting the reenlistment decisions of first-term Air Force airmen. Both types of models are constructed using 18 independent variables to capture the economic and Air Force policy conditions at the time each airman made a reenlistment decision. The variables included pecuniary factors, demographic factors, aptitude, experience, and the quarter in which the reenlistment decision was made. The models were estimated using data which covered the January 1975 through March 1982 time period, and validated the resulting models over the April 1982 to March 1986 time period data.

Each of the major Air Force Specialties (AFS's) were modeled using a separate probit equation estimated on individual level data for all airmen in an AFS eligible to make a decision during the estimation sample time frame. The resulting probit equations were used to predict the reenlistment decisions of airmen eligible to make reenlistment decisions over the validation sample time frame.

Three neural network models were created using the backpropagation learning algorithm, each with different criteria for stopping training. The first, BP Hold, computed the validation sample root mean square error (RMSE) after each training pass through the estimation sample data. Training was stopped when the RMSE was minimized. The other two models, BP Tri-sample and BP Temporal split the original estimation sample into a pre-estimation sample and a pre-validation sample. The BP Tri-sample model randomly split the original estimation sample into the two subsamples, while the BP temporal model split the samples so that they covered two separate time periods. For both the BP Tri-sample and the BP Temporal models training was done only on the pre-estimation sample, and testing tracked the RMSE of the pre-validation sample. When this RMSE was minimized the network was retrained on the full estimation sample, and training was stopped when the RMSE from the full estimation sample matched the RMSE from the pre-validation sample.

Wiggins and Engquist used simulation R^2 to measure the performance of each model's predictions. An R^2 of one implies a perfect fit whereas a zero implies a model which performs no better than the in-sample mean.

$$R^2=1-\frac{\sum (Predicted_i-Actual_i)^2}{\sum (ActualMean-Actual_i)^2}$$

The validation sample results of the neural networks compared to the probit models are shown in Table 3. None of the simulated R^2 were very high, and all of the models had very low explanatory power, as is often the case with individual level data. In virtually all cases the neural network models performed better than the probit models currently in use.

TABLE 3: VALIDATION SAMPLE RESULTS

AFS Network	Simulation R^2 by Modeling Technique			
	Probit	BP Hold	BP Tri-Sample	BP Temporal
Air Traffic Control	.139	.222	.154	.205
Missile System Maintenance	-.194	.116	-.173	-.035
Jet Engine Mechanic	.269	.368	.141	.365
Communications Electronics	.155	.244	.241	.316
Vehicle Maintenance	.198	.331	.300	.312

Source: Wiggins and Engquist (1993)

C. COMPARISONS OF NEURAL NETWORKS TO CLASSICAL FORECASTING METHODS IN SELECTED CIVILIAN AREAS

Several studies have been done comparing neural networks with classical forecasting methods in areas outside of military manpower and personnel analysis. These areas include bond rating, bankruptcy prediction, and stock price prediction. These areas have some common characteristics with military forecasting areas, which allow them to be reviewed in the context of this thesis.

Surkan and Singleton [Ref. 10] compare neural networks to multivariate discriminant analysis at the task of separating two non-contiguous classes of bonds. Bond ratings have both economic significance, as higher ratings command lower interest rates, and investor interest, as investors wish to anticipate changes in interest rates due to changes in company circumstances.

For this research Surkan and Singleton collected data on the eighteen Bell Telephone operating companies divested by American Telephone and Telegraph Company (AT&T) in 1982, for the years from 1982 through 1987. They use the seven dependant variables related to leverage, coverage, and profitability which are taken into account by the major rating companies (Moody's or Standard and Poor's) when awarding bond ratings. Those variables and their definitions are shown in Table 4. In both the linear discriminant and the neural network model these seven variables were used to predict whether a bond would be assigned a highest quality (Aaa) [group one or a medium quality (Aa1, Aa2, or Aa3) [group two rating.

Linear discriminant functions are estimated using the two bond groups as dependent variables and the seven financial ratios as explanatory variables. Fifty-six observations were used in a hold-one-out approach by iteratively calculating the model over 55 observations and classifying the 56th. The discriminant models correctly predicted 12 of 30 for group one (40%), 10 of 26 for group two (38%), and 22 of 56 overall (39%).

TABLE 4: MODEL VARIABLES AND THEIR DEFINITIONS

Variable	Definition
LEVERAGE	Debt divided by total capital - a measure of the bondholders' security
COVERAGE	Pre-tax interest expense divided by income - a measure of the company's ability to pay bondholders from current income
ROE	Return on equity or income - a profitability measure
CV of ROE	Coefficient of variation of ROE calculated over the past five years - an indication of the stability of profitability
TA	Logarithm of the total assets - a measure of size
FLOW	Construction costs divided by total cash inflow - a measure of the capacity for func construction costs without increased borrowing
TOLL	Toll revenue ratio - an indication of the effect of divestiture on profitability

Source: Surkan and Singleton (1990)

Three neural network models were created for this analysis. All three models used backpropagation as the model architecture, with seven input neurons and two output neurons, one for each input or output variable. Model one used one hidden layer with 14 neurons in that layer, while models two and three used two hidden layers. Model two used five and ten neurons in its respective hidden layers, while model three was constructed with ten and five neurons in the two hidden layers. The 56 observations used to build the discriminant analysis model were used to train the three neural network models. These neural network models were then tested on a holdout sample of 20 observations each, for group one and group two data, previously unknown to the neural network models. Results for both the neural network models and the discriminant analysis model are shown in Table 5.

As shown in Table 5, neural network models significantly out-performed linear discriminant models in all cases. A shortcoming with this study is that no forecasts were made on the holdout sample (40 observations) with the linear discriminant model. A better test would have built a single linear discriminant model with the first 56 observations and tested the model on both the holdout sample and the model building sample. This would have allowed a direct comparison of the neural network models with the linear discriminant model over a sample new to each model.

Odom and Sharda [Ref. 11] compare neural networks to multivariate discriminant analysis at the task of bankruptcy risk prediction. Failure analysis of banking firms using financial ratios are used by management, prospective investors, and auditors. Ratio analysis is the most common technique used to predict whether or not an institution will become bankrupt.

Bankruptcy prediction is most commonly done using discriminant analysis of five financial ratios obtained from accounting data.[2] For this study data were obtained from

---

[2]These ratios are:
1. Working Capital/Total Assets
2. Retained Earnings/Total Assets
3. Earnings before Interest and Taxes/Total Assets

Moody's Industriales Manuals on 129 firms. The sample consisted of 65 bankrupt and 64 nonbankrupt firms. This sample was further split into two subsamples, a training set of 38 bankrupt and 36 nonbankrupt firms, and a testing set of 27 bankrupt and 28 nonbankrupt firms.

TABLE 5: CLASSIFICATION ACCURACY RESULTS

Network	NN-7,14,2	NN-7,5,10,2	NN-7,10,5,2	Linear Analysis
Bond Class	Training Sample (56 Observations)			
Highest	(27) [90%	(28) [93%	(30) [100%	(12) [40%
Medium	(15) [58%	(20) [77%	(21) [81%	(10) [38%
Both	(42) [75%	(48) [86%	(51) [91%	(22) [39%
	Testing Sample (40 Observations)			
Highest	(17) [85%	(18) [90%	(20) [100%	No Test
Medium	(9) [45%	(14) [70%	(15) [75%	No Test
Both	(26) [65%	(32) [80%	(35) [88%	No Test

Source: Surkan and Singleton (1990)
Note: Table entries give (number) and [percent correctly classified

One discriminant analysis and one neural network model were created for this study. SAS DISCRIM was the program used for the discriminant analysis model. The neural network model used backpropagation as the network architecture, with five

4. Market Value of Equity/Total Debt
5. Sales/Total Assets [Ref. 7 p. II-164]

input neurons, five hidden neurons in one hidden layer, and one output neuron. To examine the robustness of both types of models, three separate groups of training data were used on both models. The first used all of the data available in the training subset of 38 bankrupt and 36 nonbankrupt firms, referred to as the 50/50 training set. The training data set was then randomly adjusted to be more realistic of the real world ratio of nonbankrupt firms to bankrupt firms. The second subsample consisted of 36 nonbankrupt to nine bankrupt firms, while the third subsample consisted of 36 nonbankrupt to four bankrupt firms. These are referred to as the 80/20 and the 90/10 training sets. Essentially, one discriminant analysis and one neural network model was created on each training set of data, then tested on the holdout sample.

The results of the tests of the models on the holdout sample are shown in Table 6. The neural network models clearly outperformed the discriminant analysis model in the task of bankruptcy prediction. The neural network model predicted 81.48 percent of the bankrupt firms compared to 59.26 percent for the discriminant analysis model based on the 50/50 training sample, 77.78 percent to 70.37 percent based on the 80/20 sample, and 77.78 percent to 59.26 percent based on the 90/10 sample.

At the task of correctly predicting nonbankrupt firms, the results were mixed. For the 50/50 training sample models the discriminant analysis model correctly predicted 89.29 percent to the neural network model's correct rate of 82.14 percent. The discriminant analysis model also outperformed the neural network model based on the 80/20 sample by predicting 85.71 percent to 78.57 percent. However the neural network model outperformed the discriminant model based on the 90/10 training sample by correctly predicting 85.71 percent compared to 78.57 percent for the discriminant model.

TABLE 6: COMPARISON OF DISCRIMINANT ANALYSIS AND NEURAL NETWORK MODELS ON THE HOLDOUT SAMPLE

Training sample proportion	Neural Network	Discriminant Analysis
Bankruptcy Prediction (27 observations)		
50/50	(22) [81.18%	(16) [59.26%
Medium	(21) [77.78%	(19) [70.37%
Both	(21) [77.78%	(16) [59.26%
Nonbanckruptcy prediction (28 observations)		
50/50	(23) [82.14%	(25) [89.29%
80/20	(22) [78.57%	(24) [85.71%
90/10	(24) [85.71%	(22) [78.57%

Source: Odom and Sharda (1990)
Note: Table entries give (number) and [percent correctly classified

The results of this study indicate that neural networks have promise for prediction purposes in the area of bankruptcy analysis. The neural networks significantly outperformed the discriminant analysis model for bankruptcy prediction, and performed better at nonbankruptcy prediction as the ratio of bankrupt to nonbankrupt firms declined in the training sample. However, discriminant analysis has several shortcomings which could lead to neural networks appearing favorably in this comparison. Afifi and Clark [Ref. 12] list the following as possible trouble areas for discriminant analysis:

1. A simple random sample from each population is assumed. As this is often not feasible, the sample taken should be examined for possible bias errors.

2. If some of the variables are dichotomous and one of the outcomes rarely occurs, then logistic regression analysis should be considered as a modeling technique rather than discriminant analysis.

Possible ways to improve this study would be to use more than the five ratios as inputs to the models, and to use multiple hidden layered neural networks with various numbers of neurons in those hidden layers.

Yoom and Swales [Ref. 13] compared the predictive power of a neural network model with that of a multiple discriminant analysis model at the task of predicting stock price performance. Both qualitative and quantitative variables help form the basis of investor stock price expectations and influence investment decision making. These variables also form the basis of stock price fluctuation; if investors believe that a company has the potential for strong growth, demand for the stock will rise as will the price. Conversely, if investors feel that a company is weak financially, demand for its stock will decrease and drive down the price. Thus a model predicting stock price performance should contain those variables, both quantitative and qualitative, that influence investor decision-making.

Yoom and Swales reviewed previous studies in which multiple discriminant analysis models were used to predict stock price performance. These studies utilized quantitative financial variables to construct their models, which have reasonably good predictive results. These models provide the basis for Yoom and Swales' models. In addition, they use qualitative variables gleaned from companies' annual reports. Content analysis was done on the presidents' letters to shareholders of the companies included in this study. The most important recurring themes of these reports are analyzed for frequency and percentage of the report, and used as inputs to both the multiple discriminant analysis and the neural network models.

The data for this study are taken from the Fortune 500 and Business Week's "Top 1000." These sources provide the quantitative variables used by investors, while the president's letters to investors are used to determine which qualities are important to the individual companies.

The Fortune 500 sample includes observations on the 58 firms from the five industries that offer investors the highest total return in the year of the report. The Business Week sample includes observations from the 40 firms in the 10 industries that are reported to have offered the highest total return to investors. Both samples were subdivided into two groups; group one consisted of those firms with the highest market valuations for their industry, while group two consists of those firms with the lowest market valuations. A multiple discriminant analysis model was then constructed, including both the quantitative and qualitative variables previously discussed, and the model was derived from the Fortune 500 sample. The output parameters for the model are whether the firm is a well-performing or a poor-performing firm.

A neural network model was also created using the data form the Fortune 500 sample. The model used backpropagation as the network architecture, with two hidden layers containing four neurons in the first and one neuron in the second hidden layer. The network used one output neuron. Both the neural network and the multiple discriminant analysis models were then tested on the Business Week sample.

The results of both the tests on the training data and the testing data are shown in Table 7. On the training set data (Fortune 500 sample) the multiple discriminant analysis model correctly classified 21 of 29 companies into group one, and 22 of 29 companies into group two. On the testing set (Business Week sample) the multiple discriminant model correctly classified 14 of 20 into group one, and 12 of 20 into group two.

The neural network model performs significantly better than the multiple discriminant analysis model. The neural network model correctly classified 25 of 29 firms into group one and 28 of 29 firms into group two on the training data. For the testing data set the model correctly classified 18 of 20 companies into group one and 13 of 20 companies into group two.

TABLE 7: PERFORMANCE OF THE MULTIPLE DISCRIMINANT ANALYSIS MODEL AND THE NEURAL NETWORK MODEL ON THE TRAINING AND TESTING DATA

Group	Neural Network	Discriminant Analysis
Training Data (58 observations)		
Group 1	(25) [86%	(21) [72%
Group 2	(29) [96%	(22) [76%
Mean	[91%	[74%
Testing Data (40 observations)		
Group 1	(18) [90%	(14) [70%
Group 2	(13) [65%	(12) [60%
Mean	[77.5%	[65%

Source: Yoom and Swales (1990)
Note: Table entries give (number) and [percent correctly classified

D. NEURAL NETWORKS FOR TIME SERIES FORECASTING

Hill, O'Conner, and Remus [Ref. 14] evaluated neural network models for time series forecasting. They compared neural network models with three classes of traditional time series forecasting models: statistical methods, human judgement methods, and naive-forecasting methods. Hill et al. compared neural networks with models from each class of traditional model in side-by-side experiments over the same data sets. The comparisons were done on monthly, quarterly and yearly time series data.

The data for the comparisons came from the "M-competition," described by Hill et al. as 1001 real time series gathered by Makridakis. These time series were gathered for a competition in which various groups of forecasters were given all but the most recent data points in a systematic sample of 111 of the series. The forecasters, all experts in their area of forecasting, were then asked to make time series forecasts for the most recent points in the 111 series. Each competitor's forecasts were then compared to the actual values in the holdout samples.

In the original "M-competition' 24 different forecasting methods were used. Hill et al. chose six methods which performed relatively well in the competition, out of the set of 24 from which to compare neural network models. From the statistical method category three models were chosen: the deseasonalized simple exponential smoothing, the Box-Jenkins, and the deseasonalized Holt exponential smoothing method. From the human judgement-based methods the authors chose graphical forecasts and a combination model.[3] The authors also included a naive forecasting model in which next period's forecast is whatever happened in the prior period.

Two neural network models were formulated. The first (NN-I) forecast all periods in the forecast horizon simultaneously. The second neural network model (NN-2) forecast for the first period of the forecast horizon, then fed that forecast back into the network as input to forecast into the second period of the forecast horizon, and so on. The authors used the first two time series from each of the three categories (monthly, quarterly, and annually) of time series data sets to develop the structure of the two neural network models. These series were omitted from the analysis, leaving 105 series in total (18 annual, 21 quarterly and 66 monthly). Upon further investigation, one monthly series (series 106) was found to have three major discontinuities, and was eliminated from the monthly database. Forecast accuracy was compared on the basis of absolute percentage forecast error (APE).[4] Because the forecasts were not

---

[3]This model is the average of the forecasts of six statistical methods (deseasonalized single exponential smoothing, deseasonalized adaptive response rate exponential smoothing, deseasonalized Holt's exponential smoothing, deseasonalized Brown's linear exponential smoothing, Holt-Winter's linear and exponential smoothing, and Carboni-Longini filter method).

[4]APE = $(1/N)(\mathrm{Sum}|E_t/X_t|)*100$

where: N = Number of residuals
X_t = Actual value of forecast
E_t = Predicted value of forecast t - X_t

statistically independent nor necessarily normally distributed, the APE's of the neural network models were compared with the traditional model forecasts using the paired t-test.

The second type of neural network model (NN-II) was found to provide a higher accuracy than the first type (NN-I). Given the overall superiority of NN-II, the authors focused on it when comparing the neural network model with the traditional models. Table 8 presents the mean absolute percentage errors (APE's) and their standard deviations for both the neural network models and the traditional models for the annual, quarterly, and monthly restricted data sets.

Table 8 shows mixed performance results for the neural network model on the annual time series compared to the traditional models. The neural network model performed significantly better than the deseasonalized exponential smoothing and the naive models, but significantly worse than the human judgement models using the graphical method and the six methods combined.

On the quarterly and monthly time series data the neural network model performed significantly better than the traditional forecasting methods. In only one case (deseasonalized exponential smoothing over the monthly time series) did the neural network not clearly outperform the traditional models, and in that case the neural network model performed at least as well as the traditional model.

TABLE 8: COMPARISON OF A NEURAL NETWORK MODEL WITH TRADITIONAL MODELS FOR TIME SERIES FORECASTING

Network	Annual	Quarterly	Monthly Restricted
NN-2	14.2 (17.1)	15.3 (17.1)	13.6 (14.3)
Deseasonalized Exponential Smoothing	15.9 (17.0) **	18.7 (27.0) **	15.2 (33.1)
Box-Jenkins	15.7 (22.8)	20.6 (40.8) *	16.4 (26.9) ***
Deseasonalized Holt's	12.1 (16.0)	26.9 (50.2) ***	19.2 (47.5) ***
Graphical Human Judgment	12.5 (12.5) **	20.5 (34.5) **	16.3 (22.8) ***
Six Methods Combined	12.6 (16.1) *	21.2 (38.3) **	16.7 (41.0) ***
Naive	16.4 (16.7) ***	20.0 (27.8) ***	27.0 (40.4) ***

Source: Hill et al. (1990)
Note: Table entries give Mean (and Standard Deviations) of APE's for each method across each series grouping
Results of comparison paired t-tests with NN-II are shown for * for .05, ** for .01, and *** for .001 levels.

The authors of the study conclude that neural networks as predictors for time series forecasting show great promise. However, they caution that finding the best neural network structure to learn the underlying functional form of the data set is a formidable task

Wiggins and Engquist [Ref. 9] examined the use of neural network as modeling tools for the Air Force personnel system. On an aggregate level the Air Force personnel system has three major flow rates: non-prior service accessions (NPS), prior service accessions (PS), and separations. Currently only voluntary separations are modeled using the reenlistment rates for first term (RELRT1) and second term (RELRT2) airmen. Wiggins and Engquist compare the predictive power of three neural network models with those of two more traditional modeling techniques for predicting Air Force personnel flows.

Traditionally Air Force personnel flows have been modeled using ordinary least squares (OLS) to separately estimate each flow rate equation and generalized least squares (GLS) to simultaneously estimate the four (NPS, PS, RELRT1, and RELRT2) flows. Wiggins and Engquist estimate the equations using data over one time period, October 1979 through September 1987, and validated their performance over the time period October 1987 through September 1988.

Wiggins and Engquist created three neural network models, using stopping criteria similar to those used in their individual reenlistment model, described earlier in this chapter. The BP Hold method stopped training when performance was best on the actual validation sample. The BP Temporal method terminated training when performance was best on a temporal hold out sample. The third training heuristic stopped training when the second derivative of the in-sample

RMSE with respect to the amount of training, switched from negative to positive for the second time. This network was designated the BP Inflection network.

A comparison of the performance of the three neural network models and the two regression techniques, on the validation sample, is shown in Table 9. The R^2 value for comparison is the same comparison statistic described earlier in the chapter for the Wiggins and Engquist article.

The authors noted that in nearly all cases the neural network models clearly outperformed the traditional regression models. In several cases the neural network models explained more than twice the out-of-sample variations when compared to the OLS or GLS models.

TABLE 9: VALIDATION SAMPLE RESULTS

Modeling Technique	Simulation R^2			
	NPS	PS	RELRT1	RELRT2
OLS	.618	.378	.288	.569
GLS	.606	.317	.237	.323
BP Temporal	.487	.633	.683	.736
BP Hold	.647	.633	.774	.736
BP Inflection	.644	.550	.772	.436

Source: Wiggins and Engquist (1993)

E. CONCLUSION

The articles reviewed in this chapter show that neural networks hold promise as alternatives to more traditional forms of modeling. The remainder of this thesis is an exploration of the use of neural networks to a problem specific to military manpower analysis, namely, that of predicting reenlistment.

IV. DATA AND METHODOLOGY

A. INTRODUCTION

Determining the efficacy of neural network models for military manpower and personnel analysis, requires tests that compare the results and outcomes of both neural networks and traditional data analysis techniques using the same data. Traditional data analysis techniques based on accepted econometric principles should be used for a baseline model, against which neural network models can be compared. This type of comparison is essential to assess how neural network models can perform as tools for the military manpower and personnel analyst.

Features of the assessment of a neural network model for this thesis follow:

1. Acquire a large manpower data set for which a standard regression model has been developed.

2. Randomly subset the data into a training data set and a testing data set.

3. Use the training data set to estimate a traditional data analysis model, based on accepted econometric techniques.

4. Develop two neural network models using NeuralWare software: (i). Neural network model one using the training data set with the same variables used to develop the traditional data analysis model. (ii). Neural network model two using the training data set with an expanded number of variables.

5. Apply both the neural network models and the traditional data analysis model to the testing data set, to test the predictive power of the models.

6. Evaluate the results of the tests, compare the outputs of the models, and make recommendations based on those comparisons. The criterion used for comparisons of the models is the number correctly predicted on the testing data set.

The remainder of this chapter describes the data set used for this thesis, the variables selected to build the models, and the methodology used to develop both the traditional data analysis model and the neural network models.

B. DATA

The data used for this thesis were extracted primarily from the 1985 DoD Survey of Officer and Enlisted Personnel [Ref. 15]. The 1985 survey has been matched by social security number with personnel records to obtain information on respondents' military status in 1989.

The 1985 survey was conducted by the Defense Manpower Data Center (DMDC) to provide information for the services to help improve force readiness and retention. The survey was conducted in response to a mandate by the Deputy Secretary of Defense for Force Management and Personnel, with an emphasis placed on military families, who were recognized as extremely important to the retention and readiness of the services.

Table 10 describes the nine sections of the survey. The population from which the survey was drawn consisted of active duty officers and enlisted members worldwide who were on active duty as of 30 September 1984. Members considered new accessions, those with less than four months active duty service, were excluded from the population. The survey was administered to approximately 132,000 active duty military members, providing a large cross-sectional sample of the U.S. military.

TABLE 10: THE 1985 DOD SURVEY OF OFFICERS AND ENLISTED PERSONNEL TOPIC AREAS

Section	Questionnaire Topic Area
1	Military Information--Service, Paygrade, military occupation, term of enlistment
2	Present and Past Locations--length of stay, expected stay, and problems encountered at present and past duty stations
3	Reenlistment/Career Intent--expected years of service, expected rank when leaving the service, and probable reenlistment behavior
4	Individual and Family Characteristics--basic demographics such as age, sex, and marital status
5	Dependents--basic demographics from Section 4, and whether or not dependents were handicapped
6	Military Compensation, Benefits, and Programs--benefits received for military service, and availability and satisfaction with family programs
7	Civilian Labor Force Experience--members' civilian work experience and previous earnings
8	Family Resources--household's civilian work experience and earnings, and non-wage or salary sources of earnings
9	Military Life--satisfaction with various aspects of military life, including pay and allowances, interpersonal environment, and benefits

Source: 1985 DoD Survey of Officers and Enlisted Personnel

This thesis compares neural network models and a more traditional model in analyzing the re-enlistment decisions of a relatively homogeneous group of service members. The sample chosen for this comparison includes male, Navy enlisted personnel, with 24 to 72 months of active duty service. To ensure that all members of the data set were afforded an opportunity to make a re-enlistment decision prior to the 1989 status variable being matched with the survey data, only those members who were within three years of their end of obligated service were included. To avoid the effects of atypical enlisted personnel, the sample was further constrained to personnel in the paygrades E-3 to E-6, who were 30 years of age or younger when they first enlisted in the military. Finally, those observations which contained missing or unrealistic values were also omitted from the sample data set. The sample size was 680 observations.

C. VARIABLE DEFINITIONS

Variables expected to affect the reenlistment decision were chosen based upon a logistic regression model developed and estimated by Kathy Kocher and George Thomas at the U.S. Navy Postgraduate School, Monterey, California. The following variables will be used to develop the traditional data analysis model and neural network model one. The variables which will be used to develop neural network model two will consist of all the following variables, and the variables discussed in section D of this chapter, and described in Table 12.

1. Dependent Variable (STATUS)

The dependent variable STATUS is a dichotomous variable measuring the actual reenlistment behavior of the sample members. The variable is equal to one if the individual remained on active duty three years after the survey, and equal to zero if he separated by that time.

2. Independent Variables

The independent variables chosen for this analysis fall into one of five general categories: Demographics, Military characteristics, Educational level, Level of perceived employability and Satisfaction with Military Life and Military Benefits.

a. Demographic Variables

(1) Age Upon Entering Active Duty Status ENTRYAGE is the member's age when he entered active duty in the Navy. ENTRYAGE is computed by subtracting the amount of time the member has served on active duty from his reported age at the time of the survey. As a member's age at entering active duty goes up, the time remaining in his work career decreases, giving him less time to establish a second career. Therefore, ENTRYAGE is hypothesized to have a positive effect on the probability of reenlistment.

(2) Race Race is measured using the three dummy variables, WHITEOTH, BLACK, and HISPANIC. A dummy variable is coded as a one if the member falls into that category, and as a zero if he does not fall into that category. Past studies have shown that minorities reenlist at a higher rate than caucasians, possible due to perceived lower employment opportunities for minorities in the civilian labor market. Minorities other than people of African American or Hispanic descent are categorized with caucasians in the category WHITEOTH to keep the number of categories low and ease the modeling problem.

(3) Family and Marital Status Family and Marital Status is categorized by the four dummy variables Single No Children (SNC), Single With Children (SWC), Married No Children (MNC), and Married With Children (MWC). The category into which the member fell was coded as a one, while those categories in which he did not fall were as a coded zero. As a member takes on more responsibility and dependents, his ability to change careers decreases. This leads to the hypothesis that the categories SWC, MNC, and MWC will have a positive effect on the probability of reenlistment, compared to the base category of SNC.

b. Military Characteristics

(1) Rank A member's rank is measured using three dummy variables: E3, E4, and E5/6. The E5 and E6 paygrades are combined because members in those ranks are normally beyond their first enlistment and will exhibit many of the same reenlistment behaviors. Increased rank leads to increased pay and benefits, decreasing the incentive to leave the military for higher paying civilian opportunities. Rank is then hypothesized to have a positive effect on the probability of reenlistment.

(2) Military Occupation A member's military occupation is recoded into the dummy variable, Technical Occupation (TECOCC). If a member's military occupation fell into the electronic equipment repair, the communications and intelligence, the medical and dental, or other technical fields, then TECOCC was coded as a one. If the member's military occupation fell into direct combat, support and administrative, electrical/mechanical equipment repair, crafts, service and supply, or a non-occupational field, then TECOCC was coded as a zero. Those members with a technical occupation have skills that are valuable in the civilian work force, and therefore, a member who falls into the TECOCC category should have a decreased probability of reenlistment, compared with a member who does not have a technical occupation.

c. Education Level

A member's educational level was recoded into the dummy variables of having a high school degree (HSDEG) or having some type of high school certificate (HSCERT). If a member graduated and received a high school diploma then he fell into the category of having a high school degree and HSDEG was coded a one and HSCERT was coded a zero. If a member received a GED certificate, a high school completion/attendance certificate, or a home study diploma, then he fell into the category of having a high school certificate and HSCERT was coded a one and HSDEG was coded a zero. Those members who had no certificate or diploma were dropped from the data set. Those members who do not have a high school diploma should have reduced chances for a perceived "good" job in the civilian labor market. Therefore, not having a high school diploma should increase the probability of reenlistment.

d. Level of Perceived Employability

A major factor in whether a member decides to reenlist or not is his perceived chances of finding a good civilian job. In the original DoD Survey, a member was asked to rate, on a scale of one to ten, what he felt his chances were of being able to get a good civilian job if he left the military at the time of the survey. This response was recoded to a dummy variable CIVJOB, receiving a one if the member responded to the original question with an answer of seven or higher, and a zero if he felt his chances of getting a good civilian job were six or less.

e. Satisfaction with Military Lifestyle and Military Benefits

A major portion of the 1985 DoD Survey deals with the member's satisfaction with military life and benefits of being in the military. However, correlation analysis shows that those satisfaction variables that have high predictive power for the reenlistment decision also are highly correlated with each other. Although multicollinearity will have little effect on the overall fit of a model, and thus little effect on the use of that model for prediction or forecasting, the variances of the variables will increase and the computed t-scores will fall. This rise in variances and fall in t-scores will reduce the explanatory power of the traditional data analysis model.

One solution to the problem of multicollinearity between independent variables is factor analysis. Factor analysis will yield explanatory variables which are uncorrelated and thus do not reduce the explanatory power of the traditional model. For this reason, factor analysis was undertaken using the satisfaction variables to compute two new variables, FACTOR1 and FACTOR2. Table 11 shows the rotated factor pattern scores for the satisfaction variables included in the analysis. As satisfaction with the military lifestyle and military benefits increase, the probability of reenlistment should also increase, all other variables held constant. An increase in a satisfaction variable will have a positive increase in either FACTOR1 or FACTOR2, which will lead to an increase in the probability of reenlistment.

(1) FACTOR1 FACTOR1 loads heavily on the satisfaction with military lifestyle variables. Those variables include: job satisfaction, satisfaction with working conditions, satisfaction with job training, satisfaction with job stability, satisfaction with a member's co-workers, satisfaction with job security, satisfaction with personal freedom, satisfaction with promotion opportunity, satisfaction with the opportunity to serve his country, satisfaction with personal friendships, and satisfaction with military moves and moving frequency.

(2) FACTOR2 FACTOR2 is loaded heavily on the satisfaction with military benefits variables. Those variables include: satisfaction with medical care, satisfaction with dental care, satisfaction with commissary services, satisfaction with future retirement benefits, satisfaction with military pay, and satisfaction with Veterans Educational Assistance Program (VEAP) benefits. Satisfaction with the military family environment loads heavily on FACTOR2.

Satisfaction with military pay loads more heavily on FACTOR2 than on FACTOR1 but the loading is relatively close.

TABLE 11: ROTATED FACTOR PATTERN SCORES

Satisfaction Variables	FACTOR1	FACTOR2
Overall Job	0.71266	.
Work Conditions	0.62253	.
Job Training	0.55176	.
Job Stability	0.54597	.
Co-Workers	0.51141	.
Job Security	0.49178	..
Promotions	0.47001	.
Personal Freedom	0.46376	.
Ability to Serve Country	0.42604	.
Family Environment	0.41481	0.37981
Friendships	0.36824	.
Moving	0.35458	.
Medical Care	.	0.76467
Dental Care	.	0.69765
Commissary Services	.	0.50460
Retirement Benefits	.	0.43947
Pay	0.38413	0.43609
VEAP Benefits	.	0.41571

Note: Values less than 0.3 have been printed as '.'

D. METHODOLOGY

1. Traditional Data Analysis Model

Multivariate data analysis is used to quantify the relationship between the dependent variable STATUS, and the independent or explanatory variables discussed earlier in this chapter. The estimation technique used here is binomial logistic regression, suitable for the analysis of a dichotomous dependent variable such as STATUS.

The model is based on the cumulative logistic distribution function, and has the following functional form:

$$Ln(P_i/1-P_i) = \beta_0+\beta_1X_{1i}+\beta_2X_{2i}+...+\beta_nX_{ni}+\epsilon_i$$

The estimated value P_i is interpreted as the probability that member i will reenlist for active duty, given his set of explanatory variables $(X_1, X_2, ..., X_n)$. The β_n's represent the estimated coefficients associated with the respective X_n's. β_0 is the constant term, and ϵ is the stochastic error term.

2. Neural Network Models

The neural network models will be constructed using NeuralWare, a commercially available brand of neural network software. It was chosen for use in this thesis because it was readily available at the Naval PostGraduate School.

Construction of a neural network model is often considered an art rather than a hard science. For this reason, the methodology of creating a neural network model may seem rather haphazard. Both neural network models will be constructed using the backpropagation learning algorithm with the generalized delta rule. The Tan H transfer function will be used as the initial transfer function because the networks are concerned with prediction as their basic feature. Neural network model one will be created with the same set of variables used in the logistic regression model. Initially neural network model one will be constructed using NeuralWare's default settings for learning rate (alpha) and momentum. The neural network model one will initially be constructed with a single hidden layer containing five neurons, and will be trained for 500,000 learning cases. Epoch size, or the number of training cases the network looks at before it updates itself, will be changed from the default setting of 16 to a factor of the data set size, 68. Learning transition point, the point at which the network begins to decrease the learning rate to prevent oscillations in the network as it attempts to move down the error structure, will be moved from 10,000 to 50,000 iterations to allow the network more time to train at each training rate. NeuralWare recommends that the learning transition be increased as the size of the data set increases.

Subsequent variations of neural network model one will be constructed using varying numbers of neurons in up to two hidden layers. The model chosen as the final neural network model one will be the model that has the best predictive ability on the holdout testing data set.

In order to test the ability of a neural network to model a problem that a researcher is unfamiliar with or that has no apparent underlying theoretical model, a second neural network model will be constructed and compared to neural network model one. Neural network model two will be constructed using an extended data set that includes all of the theoretically sound variables used to develop neural network model one, plus the variables shown in Table 12. Some of the variables shown in Table 12 are theoretically sound for predicting reenlistment, while others such as MILHOUR are merely noise that the neural network should be able to ignore. Neural network model two will be constructed using the same architecture as neural network model one, and the emphasis of the comparison will be whether or not the two neural network models have comparable partial effects of explanatory variables on reenlistment.

In summary, this thesis will make two comparisons. First, it will compare the results of a neural network model (neural network model one) to a traditional econometric data analysis method (logistic regression) for predicting reenlistment in the Navy. These two models will be constructed using the same data set and the same set of variables. A second neural network model will also be developed (neural network model two), but using an extended set of variables on the same data set as the first two models. A comparison will then be made between the two neural network models to determine if there are significant differences between the two neural network models. The following chapter describes the logistic regression model and its results.

TABLE 12: EXTENDED DATA SET VARIABLES FOR THE CONSTRUCTION OF NEURAL NETWORK MODEL TWO

Variable	Description of the Variable
SPACTIVE	A dummy variable coded "1" if the member had a spouse on active duty in the military, and "0" otherwise
SEATIME	Months of career sea time
OSEATIME	Months of career oversea's time
INCOME	Total family income
PCSMOVE	Number of permanent change of station moves a member had made during his career
MOMSED	Total years of a members mothers education
OFDTYJOB	Number of weekly hours spent on an off duty job
CIVJOBOF	A dummy variable coded "1" if a member had ever received a "good" civilian job offer, and "0" otherwise
MILHOUR	Military hour that the member was surveyed
NUMENLST	Number of enlistment when the member was surveyed
DEBT	A categorical variable, between one and seven, of a members total household debt

Source: 1985 DoD Survey of Officers and Enlisted Personnel

V. RESULTS OF THE LOGISTIC REGRESSION MODEL

A. DESCRIPTIVE STATISTICS

Table 13 displays the means, standard deviations, and ranges for the variables included in the final logit model. The mean values of the categorical variables can be interpreted as the percentage of the data set that hold that characteristic. For example, 12.21 percent of the data set is of African-American descent, and fall into the category BLACK. Of those members in the sample, 31.18 percent hold a technical occupation. Rank is divided into 20.73 percent E3, 38.68 percent E4, and 40.59 percent E5/6.

B. RESULTS OF THE LOGISTIC MODEL

The generally accepted criteria for assessing the overall fit of a logistic model is the -2 Log Likelihood statistic (-2 Log L). The -2 Log L has a chi-square distribution under the null hypothesis that all the explanatory variable parameters in the model are zero. The -2 Log L for the reenlistment model is computed to be 83.709 with 13 degrees of freedom. Using the chi-square distribution, the probability that the null hypothesis is true for the reenlistment model is less than .0001 (p=.0001).

TABLE 13: SIMPLE STATISTICS FOR EXPLANATORY VARIABLES IN THE LOGISTIC MODEL

Variable	Mean	Standard Deviation	Minimum	Maximum
CIVJOB	0.8118	0.3912	0	1
ENTRYAGE	19.2558	2.1955	16.00	29.83
E4	0.3868	0.4874	0	1
E56	0.4059	0.4914	0	1
BLACK	0.1221	0.3276	0	1
HISP	0.0824	0.2751	0	1
SWC	0.0176	0.1318	0	1
MNC	0.1765	0.3815	0	1
MWC	0.2000	0.4003	0	1
TECOCC	0.3118	0.4636	0	1
HSCERT	0.1618	0.3685	0	1
FACTOR1	0.0097	0.8827	-2.5631	2.0796
FACTOR2	0.0091	0.8632	-2.8052	2.4067

The results of the logit analysis of the reenlistment model are shown in Table 14. The probability of a member reenlisting in the navy is derived from the equation

$P = 1 / (1+e^{-Z})$, where

$Z = -2.15 + -.659(CIVJOB) + .045(ENTRYAGE) + .654(E4) + 1.003 (E56) + .699(BLACK) - .091(HISP) + .247(SWC) + .820(MNC) + .836(MWC) + .241(TECOCC) + .240(HSCERT) + .321(FACTOR1) + .181(FACTOR2).$

TABLE 14: RESULTS OF THE LOGISTIC REGRESSION REENLISTMENT MODEL

Variable	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square
INTERCEPT	-2.1550	0.7747	7.7016	0.0055
CIVJOB	-0.6590	0.2141	9.4738	0.0021
ENTRYAGE	0.0450	0.0383	1.3829	0.2396
E4	0.6535	0.2560	6.5178	0.0107
E56	1.0031	0.2614	14.7201	0.0001
BLACK	0.6995	0.2626	7.0939	0.0077
HISP	-0.0909	0.3185	0.0814	0.7754
SWC	0.2468	0.6415	0.1481	0.7004
MNC	0.8204	0.2276	12.9896	0.0003
MWC	0.8361	0.2201	14.4307	0.0001
TECOCC	0.2409	0.1868	1.6622	0.1973
HSCERT	0.2402	0.2310	1.0818	0.2983
FACTOR1	0.3209	0.1025	9.7968	0.0017
FACTOR2	0.1812	0.1053	2.9654	0.0851

C. INTERPRETING THE RESULTS OF THE REENLISTMENT MODEL

Logistic regression model results cannot be interpreted directly from the variable parameters, because of the functional form of the model. One way to interpret the results of a logistic regression model is to establish a base case. This base case represents the reference group of variables against which comparisons can be made of the impact of individual explanatory variables on retention, holding all other variables constant.

In this instance the base case is derived from the estimated logit equation using the modal values for the categorical variables and mean values for the continuous variables. The equation for the base case, using the modeled results from Table 14 follows:

$$Z = -2.15 + -.659(CIVJOB=1) + .045(ENTRYAGE=19.256) + .654(E4=0) + 1.003\ (E56=0) + .699(BLACK=0) - .091(HISP=0) + .247(SWC=0) + .820(MNC=0) + .836(MWC=0) + .241(TECOCC=0) + .240(HSCERT=0) + .321(FACTOR1=0.0097) + .181(FACTOR2=0.0091)$$

$$Z = -1.9377$$

$$P = 1 / (1+e^{-Z}),$$

$$P = 0.1259$$

Therefore, the base case individual, a white, male E-3, single with no dependents who joined the service at age 19.25 with a high school diploma, who feels that he has a strong chance of getting a good civilian job if he leaves the military, and whose satisfaction variables give him average factor scores, will have a 12.59 percent probability of reenlisting in the Navy.

The remainder of this section is an analysis of the effects of each independent, explanatory variable on the reenlistment decision, compared to the base case set of variables.

1. Demographic Variables

a. Age Upon Entering Active Duty Service

ENTRYAGE is found to have the correct hypothesized sign, that is, the older a member was when he first entered active duty status, the more likely he was to reenlist in the Navy when his commitment was over. However, ENTRYAGE is significant only at the .25 level, making it a variable that has little reliability as an explanatory variable. The effect of a one year increase in ENTRYAGE from the base case results in a 0.5 percent increase in the probability of reenlistment.

b. Race

Being an African-American minority has the correct hypothesized sign compared to the WHITEOTH base case. The effect of BLACK is both positive and significant at the 0.01 level. The effect of being African-American as opposed to falling in the WHITEOTH category for the base case individual is a 9.9 percent increase in the probability of reenlistment.

HISP has the incorrect sign as hypothesized, but is not a significant variable. Additionally, the coefficient for HISP is small compared to BLACK. The effect of being a Hispanic minority rather than WHITEOTH for the base case individual is a decrease in probability of reenlistment of 0.97 percent.

c. Family and Marital Status

The effects of being either married, having dependents, or both all have the correct sign as hypothesized compared to the base case, single with no children (SNC) individual. Although SWC is not significant, MNC and MWC are significant at the 0.01 level. The effect of SWC compared with the base case is an increase of 3.0 percent in the probability of reenlistment. The effect of MNC and MWC are respective increases in the probability of reenlistment of 12.1 and 12.4 percent.

2. Military Characteristics

a. Rank

A member's rank when surveyed is found to have the correct hypothesized sign. The more senior a member was, the higher the probability he would have of reenlistment. Both E4 and E56 were found to be significant, E4 at the 0.05 level and E56 at the 0.01 level. The effect of being an E-4 rather than an E-3 for the base case individual is a 9.1 percent increase in the probability of reenlistment. Being an E-5 or an E-6 increased the probability of reenlistment by 15.6 percent.

b. Military Occupation

TECOCC has the incorrect hypothesized sign, but is not a significant explanatory variable up to the .19 level. The effect of having a technical occupation in comparison to the base case individual who does not have a technical occupation, is an increase of 2.9 percent in the probability of reenlistment.

3. Education Level

A member's education level was found to have the correct hypothesized sign, but is not significant at the 0.10 level. A member who had less than a high school diploma would have a higher probability of reenlisting than a member who had a high school diploma. The effect of a member not having a high school diploma in comparison to the base case individual increases the probability of reenlistment by 2.9 percent.

4. Level of Perceived Employability

CIVJOB has both the correct hypothesized sign and is significant at the 0.01 level. The effect of a member feeling that he has less than a good chance at getting a good civilian job if he left the military is an increase in the probability of reenlistment of 9.2 percent. This is compared with the base case individual, who feels that he has a good chance of getting a civilian job if he left the military.

5. Satisfaction with Military Lifestyle and Military Benefits

Both FACTOR1 and FACTOR2 have the correct hypothesized sign and are significant at the 0.01 and 0.10 levels respectively. An analysis in the change from the base case individual is inappropriate for these variables because the base case individual was assumed to have average FACTOR1 and

FACTOR2 scores, which could have occurred in many ways, due to the weighting of the factor analysis. However, it will suffice to say that a one unit increase in FACTOR1 from 0.0097 to 1.0097 will increase the probability of reenlistment by 4.0 percent, while a one unit increase in FACTOR2 from 0.0091 to 1.0091 will increase the probability of reenlistment by 2.1 percent.

D. VALIDATION OF THE LOGISTIC REGRESSION MODEL

One way to validate a prediction model is to observe how the model predicts on a data set not used in building the model. In this thesis, a random subset of 100 observations was taken from the original data set prior to constructing the logistic regression model.

A 0.5 probability cutoff was used to determine the number of correct predictions for the testing data set. That is, if the model predicted a probability of below 0.5 and the actual decision was not to reenlist, then the model was assumed to make a correct prediction. Conversely if the model predicted a probability of less than 0.5 and the member actually reenlisted, then the model made an incorrect prediction. The same logic was used for predictions above 0.5.

Overall, the model predicted 71 out of 100 (71 percent) reenlistment decisions for the testing data set. It predicted 13 out of 22 (59.1 percent) of those members who reenlisted, and 58 out of 78 (75.6 percent)of those members who decided against reenlistment. The model had a false positive rate (those members who the model predicted would reenlist, but did not) of 40.9 percent, and a false negative rate (those members the model predicted would not reenlist, but who did so) of 24.6 percent.

VI. RESULTS OF THE NEURAL NETWORK MODELS

A. NEURAL NETWORK MODEL ONE DESCRIPTION

Twenty different architectures were created for neural network model one using the methodology described in Chapter IV. The models were created using various combinations of number of neurons and number of hidden layers (one or two). The initial neural network model contained five neurons in a single hidden layer, and subsequent modifications of this architecture included hidden layers with as few as one, and as many as 100 neurons in a single hidden layer. Several networks were also constructed using two hidden layers, with various combinations of number of neurons in each layer. Initially all networks used the default settings in NeuralWare for learning rate (alpha) and momentum, but these were also varied for each network architecture. Initially all architectures used the Tan H transfer function, but were modified to use the sigmoidal transfer function also.

All of the different neural network architectures constructed for neural network model one contained the same variables used to construct the logistic regression model. They contained 17 input neurons, one for each explanatory variable included in the model. Because the output variable STATUS was a dichotomous variable, taking on a output value of either one or zero, only one output neuron was used to model the reenlistment decision. All of the various model architectures were tested on the testing data set to determine which architecture was the best at predicting reenlistment.

The best model architecture at predicting reenlistment was constructed with a single hidden layer, consisting of two neurons in the hidden layer. It used the default settings in NeuralWare for learning rate and momentum, and used the Tan H transfer function. For the remainder of this thesis this architecture will be referred to as neural network model one. Figure 14 is a pictorial depiction neural network model one.

Neural Network Reenlistment Model One

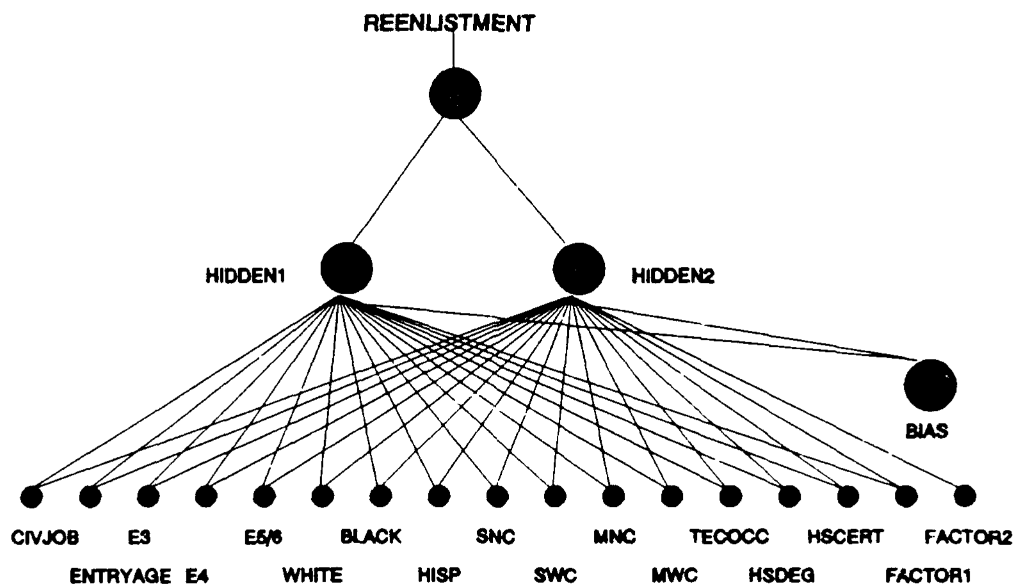


Figure 14

B. DESCRIPTIVE STATISTICS

NeuralWare provides no descriptive statistics such as mean and standard deviation of individual variables like those produced by SAS for its logistic regression package. However, a researcher can determine the range of the variables in a neural network model by entering the MinMax window in NeuralWare, where the minimum and maximum values of each variable are presented.

C. RESULTS OF THE NEURAL NETWORK MODEL ONE

NeuralWare provides no overall goodness-of-fit statistic for its model, such as the -2 Log Likelihood statistic (described in Chapter V) provided by SAS in its logistic regression output. NeuralWare also does not provide estimates of the individual variable coefficients, like the β's provided by SAS in its logistic regression package. This occurs because the nature of neural computing is a multi-step process. Inputs, in the form of explanatory variables, are submitted to the input layer of neurons. In the input layer a scaling transformation takes place so that all of the inputs have the same scale. In NeuralWare, when using the Tan H transfer function for all of the neurons in layers beyond the input layer, the transformation is linear, and the inputs take on values that range from negative one to positive one.

Once the inputs have been scaled in the input layer, the new values are sent to the first hidden layer. Here the values are weighted, summed, and run through the transfer function, in the case of this thesis the Tan H transfer function. The outputs from the neurons in the hidden layer are then sent as inputs to the output layer, where they also are weighted, summed, and run through another Tan H transfer function. The outputs are then transformed back into their original scale to determine the final output of the network for a particular set of inputs. Because of this complex nature of neural computing, no coefficient estimates such as the β's in a logistic regression equation, are produced. However, the actual weights in the individual neurons are available as an output from the network. Table 15 shows the weights that are applied to the inputs to the two neurons in the hidden layer (Hidden1 and *Hidden2*) and the weights applied to the output neuron's inputs, which come from the two hidden layer neurons and the bias neuron.

D. INTERPRETING THE RESULTS OF NEURAL NETWORK MODEL ONE FOR REENLISTMENT

The procedure for interpreting the results of an estimated neural network model is fundamentally the same as for interpreting the partial effects of a logistic regression model. A base case is first established, representing the reference values with which comparisons are made about the partial impact of individual explanatory variables on retention, holding all other variables constant.

TABLE 15: INPUT WEIGHTS FOR NEURONS IN THE HIDDEN AND OUTPUT LAYERS OF NEURAL NETWORK MODEL ONE

Input Weights for Hidden Layer Neurons

Input Neuron	Hidden 1 Weights	Hidden 2 Weights
BIAS	0.2701	-0.3170
CIVJOB	0.5934	1.3733
ENTRYAGE	-0.7471	-0.1843
E3	0.4481	2.2858
E4	0.4818	-1.2862
E56	-1.3960	-0.7249
WHITEOTH	-0.2081	0.6314
BLACK	-0.7429	-0.7972
HISP	0.7407	0.4055
SNC	0.2586	1.2269
SWC	0.2492	0.6748
MNC	-2.2040	0.6143
MWC	1.4103	-2.0984
TECOCC	1.4855	-1.6623
HSDEG	0.9805	-1.1199
HSCERT	-0.9525	1.0637
FACTOR1	-1.0944	-1.0348
FACTOR2	-0.1112	-2.9463

Input Weights for Output Layer Neurons

Input Neuron	Output Neuron Weights
BIAS	-0.1135
HIDDEN1	-0.2961
HIDDEN2	-0.4491

The same base case will be used for neural network model one as was used for the logistic regression model described in Chapter V. This will facilitate the ease of comparisons between neural network model one and the logistic regression model. Again, in this instance the base case is derived using the modal values for the categorical variables and the mean values for the continuous variables. The base case individual is a white male E-3, single with no dependents, who joined the service at age 19.25 with a high school diploma, who feels that he has a strong chance of getting a good civilian job if he leaves the military, and whose satisfaction variables give him average factor scores. Neural network model one indicates that the base case individual will have a 6.5 percent probability of reenlisting.

An important statistic, provided by traditional data analysis packages such as SAS, are those which indicate the statistical significance of the individual variables in the model. NeuralWare provides no such statistic, and therefore the impact of a unit change in an explanatory variable on the output variable (in this case STATUS) should be evaluated and considered with caution. In many cases there may be an estimated effect on retention, yet from a statistical view a null hypothesis of no effect would be supported.

The remainder of this section describes the effects on the reenlistment decision of each changing independent, explanatory variable, compared to the base case individual.

1. Demographic Variables

a. Age Upon Entering Active Duty Service

ENTRYAGE is found to have no effect on the reenlistment decision of the base case individual. That is, being an additional year older or younger when initially enlisting will have no effect on the probability of reenlistment.

b. Race

Being an African-American minority has the same sign as hypothesized compared to the WHITEOTH base case. The effect of being African-American as opposed to falling in the WHITEOTH category for the base case individual is a 0.1 percent increase in the probability of reenlistment. Being Hispanic rather than falling in the WHITEOTH category has no effect on the probability of reenlistment.

c. Family and Marital Status

The effects of being either married, having dependents, or both all have the correct sign as hypothesized, compared to the base case single with no children (SNC) individual. The effect of SWC compared with the base case is an increase of 0.1 percent in the probability of reenlistment. The effect of MNC and MWC are respective increases in the probability of reenlistment of 25.5 and 3.5 percent.

2. Military Characteristics

a. Rank

A member's rank when surveyed is found to have the correct hypothesized sign. The more senior a member was, the higher the probability he would reenlist. The effect of being an E-4 rather than an E-3 for the base case individual is a 9.5 percent increase in the probability of reenlistment. Being an E-5 or an E-6 increased the probability of reenlistment by 12.5 percent.

b. Military Occupation

Military Occupation is found to have no effect on the probability of reenlistment in the neural network model. A member with the base case characteristics will have the same probability as a member with all of the base case characteristics but has a technical military occupation.

3. Education Level

A member's education level was found to have the same sign effect as hypothesized. A member who had less than a high school diploma would have a higher probability of reenlisting than a member who had a high school diploma. The effect of a member not having a high school diploma in comparison to the base case individual increases the probability of reenlistment by 2.5 percent.

4. Level of Perceived Employability

A member's personal level of perception towards their employability has the correct hypothesized sign. A person who feels that they do not have a strong chance of finding a good civilian job if they left the military is found to have a 0.1 percent higher probability of reenlisting in the military, compared to the base case individual.

5. Satisfaction with Military Lifestyle and Military Benefits

An increase in a member's satisfaction with the military lifestyle or military benefits should result in increased reenlistment and as such, both FACTOR1 and FACTOR2 have the correct hypothesized signs. A one unit increase in either FACTOR1 or FACTOR2 resulted in an 0.1 percent increased probability of reenlistment for the base case individual. Although, because of no underlying metric, it is hard to determine the partial effects of increases in the satisfaction variables listed in Table 7, an increase in a satisfaction variable, all else held constant, will have a positive effect on the probability of a member's reenlistment.

E. VALIDATION OF THE NEURAL NETWORK MODEL ONE

Neural network model one is validated in the same way as the logistic regression model discussed in Chapter V. A 0.5 probability cutoff was used to determine the number of correct predictions for the testing data set. That is, if the model predicted a probability of below 0.5 and the actual decision was not to reenlist, then the model was assumed to make a correct prediction. Conversely if the model predicted a probability of less than 0.5 and the member actually reenlisted, then the model made an incorrect prediction. The same logic was used for predictions above 0.5

Overall, the model correctly predicted 71 out of 100 (71 percent) reenlistment decisions for the testing data set. It correctly predicted 13 out of 22 (59.1 percent) of those members who reenlisted, and 58 out of 78 (74.4 percent)of those members who decided against reenlistment. Thus the model had a false positive rate (those members who the model predicted would reenlist, but did not) of 40.9 percent, and a false negative rate (those members the model predicted would not reenlist, but who did so) of 26.6 percent.

F. NEURAL NETWORK MODEL TWO

1. Model description

Neural network model two was created using the same architecture as neural network model one, but using the extended data set described in Chapter IV. It was constructed using 28 input neurons, one for each explanatory variable in the extended data set, two hidden neurons in a single hidden layer, and one output neuron. Neural network model two had all of the same model characteristics as neural network model one regarding learning rate, momentum, learning transition point, epoch size and transfer function. The purpose behind the creation of the second neural network model was to evaluate the strength or weakness of a neural network model that has been created using a data set that contains variables that may not be theoretically sound for the problem at hand, in this case the prediction of reenlistment in the Navy. Therefore, neural network model two was constructed in the same fashion as neural network model one with the exception of using the extended data set.

Some neural network literature and researchers suggest that the "kitchen sink" approach to developing a neural network model is often appropriate [Ref. 7]. That is, if there is no apparent underlying theoretical model to begin from, or if the researcher is unfamiliar with the problem to be modeled, the network model should initially include all variables in a data set, and the neural network can determine which variables or combinations of variables will effect the output variable. In the case of this thesis, a set of variables is added to a theoretically sound set of variables to determine if the neural network model developed using the "kitchen sink" methodology (neural network model two) will resemble the model constructed using a theoretically sound base (neural network model one).

2. Model Results

Neural network model two was quite similar to neural network model one at the task of predicting reenlistment in the Navy. Neural network model two correctly predicted 72 of 100 cases in the test data set. However, as discussed in the following chapter, the partial effects of changes in the explanatory variables changed dramatically when the second model was created using the extended data set. The following chapter will also compare the results of neural network model one with the results of the logistic regression model.

VII. COMPARISON OF THE NEURAL NETWORK AND THE LOGISTIC REGRESSION MODELS

A. NEURAL NETWORK MODEL ONE AND THE LOGISTIC REGRESSION MODEL

1. Predictive Ability of Both Models

As discussed in Chapters five and six, both neural network model one and the logistic regression model correctly predicted 71 of 100 test cases. Table 16 shows that both models also correctly predicted 13 of 22 of those members who reenlisted and 58 of 78 of those members who decided to leave the military. Surprisingly, the two models did not predict the same individuals to remain with or leave the military. Of the 100 test cases, the two models predicted 90 of 100 individuals to take the same course of action. Of the individuals who the two models predicted would behave differently, neural network model one correctly predicted five of the ten cases. The logistic regression correctly predicted the five cases that the neural network model failed to predict, while incorrectly predicting the cases that the neural network model correctly predicted.

Table 16 shows that, on the training data set, both models performed comparably. Neural network model one performed slightly better overall, predicting correctly 479 of the 680 (70.44 percent) training cases, compared to the logistic regression model which predicted correctly 477 of the 680 (70.15 percent) cases. The neural network model correctly predicted 359 of the 434 (82.72 percent) members who decided not to reenlist, while the logistic regression model correctly predicted 377 (86.78 percent) of the leavers. The neural network model had a false positive rate of 51.22 percent and a false negative rate of 17.18 percent, compared to a false positive rate of 59.35 percent and a false negative rate of 13.13 percent for the logistic regression model.

TABLE 16: COMPARISON OF NEURAL NETWORK MODEL ONE AND LOGISTIC REGRESSION MODEL RESULTS

	Model			
	Neural Network	Logistic Regression	Neural Network	Logistic Regression
	Training Data Set		Testing Data Set	
Correctly Predicted	479 [70.44	477 [70.15	71 [71.00	71 [71.00
Correctly Predicted Reenlist	120 [48.78	100 [40.65	20 [60.61	20 [60.61
Correctly Predicted Leave	359 [82.72	377 [86.87	58 [86.57	58 [86.57
False Negative	[17.18	[13.13	[13.43	[13.43
False Positive	[51.22	[59.35	[39.39	[39.39
R^2	.1809	.1239	.0644	.0836

Note: Table entries give number and [percentage correctly predicted.

One possible measure for how well a model performed on the testing data set is the simulation R^2 discussed by Wiggins and Engquist, and reviewed in Chapter III of this thesis. The formula for this measure is:

$$R^2=1-\frac{\sum(Predicted_i-Actual_i)^2}{\sum(ActualMean-Actual_i)^2}$$

An R^2 of one implies a perfect fit for the data set, while an R^2 of zero would be interpreted as fitting the data no better than the in-sample mean. As is normally the case with individual level data, modeling a dichotomous outcome, both models have low R^2. The neural network model had a slightly lower R^2 than the logistic regression model on the test data set. The R^2 for both the test data set and the training data set is shown in Table 16.

2. Partial Effects of Variables on Reenlistment

Table 17 shows the partial effects of individual variables on retention for both neural network model one and the logistic regression model, as discussed in Chapters five and six.

The two models were very comparable at the task of predicting who would reenlist in the Navy. If prediction is the only question a researcher is concerned with, then the neural network model clearly performed as well as did the logistic regression model. However, often a researcher is concerned with what is affecting the output variable, in this case reenlistment, as well as with predicting who will reenlist.

Table 17 shows that the two models produced different results for the partial effects of individual variables on the probability of reenlistment.

TABLE 17: COMPARISON OF THE PARTIAL EFFECTS OF INDIVIDUAL VARIABLES ON THE PROBABILITY OF REENLISTMENT, WITH RESPECT TO THE BASE CASE INDIVIDUAL, FOR THE NEURAL NETWORK AND THE LOGISTIC REGRESSION MODELS

VARIABLE	NEURAL NETWORK	LOGISTIC REGRESSION
CIVJOB	+0.1%	*** +9.2%[1]
ENTRYAGE	No Effect	+0.5%
E4	+9.5%	** +9.1%
E5/6	+12.5%	*** +15.6%
BLACK	+0.1%	*** +9.9%
HISPANIC	No effect	-1.0%
SWC	+0.1%	+3.0%
MNC	+25.5%	*** +12.1%
MWC	+3.5%	*** +12.4%
TECOCC	No effect	+2.9%
HSCERT	+2.5%	+2.9%
FACTOR1[2]	+0.1%	*** +4.0%
FACTOR2[3]	+0.1%	* +2.1%

Notes: [1] Those variables noted with * are significant at the 0.10 level, ** at the 0.05 level, and *** at the 0.01 level. [2]Satisfaction with military pay and benefits. [3]Satisfaction with the military lifestyle.

Several of the variables (CIVJOB, BLACK, MNC, MWC, FACTOR1, and FACTOR2) had partial effects which were quite different for the two models. The neural network model appears to be loading the effects on reenlistment into two variable classes, military rank and marital status. While this is not an undesirable characteristic if a researcher's only concern is the prediction of reenlistment, it is undesirable if a researcher wishes to determine policy implications from the model.

The neural network model essentially disregards the effects of FACTOR1 (satisfaction with military pay and benefits) and FACTOR2 (satisfaction with the miliary lifestyle). This is a problem because FACTOR1 and FACTOR2 are the only variables which the military can affect (although indirectly). The military can improve pay, benefits, and the military lifestyle, which should improve satisfaction in those areas, which in turn will lead to higher FACTOR1 and FACTOR2 scores. Thus, the neural network model may lead a researcher to believe that there are no policy implications associated with variation in pay and benefits or factors affecting the military lifestyle. Intuitively this appears to decrease the usefulness of the neural network model.

Another apparent inadequacy of the neural network model is its failure to assign any effect on reenlistment to the variable CIVJOB. This variable is a member's perception about the probability of getting a good civilian job if he left the military. The neural network model essentially disregards CIVJOB as having an effect on a member's probability of reenlistment. Again, intuitively this appears to limit the usefulness of the neural network model.

However, upon further examination of the results, three positive points about the neural network model should be noted. First, the variables that the neural network found to have no effect on the probability of reenlistment for a base case individual (ENTRYAGE, HISPANIC, and TECOCC), were found to be insignificant at the 0.1 level for the logistic regression model. Second, the variables that the neural network model found to have an effect on the probability of reenlistment, had the same sign effect as in the logistic regression model. Third, several of the variables in the neural network model had partial effects which were quite close in size to their counterparts in the logistic regression model (E4, E5/6, HSCERT).

B. NEURAL NETWORK MODELS ONE AND TWO

As was discussed in Chapter VI, the predictive ability of the neural network models was quite similar. By increasing the number of variables by more than 50 percent (from 17 to 28 variables), neural network model two was able to correctly predict one case more out of the 100 case testing data set than did neural network model one. However the partial effects of the independent variables that occurs when the model is constructed on the expanded data set is disturbing. Table 18 shows the partial effects on reenlistment of a change in an explanatory variable for the base case individual for neural network models one and two. The base case individual is the same for both models for the first 17 variables; the base case for the extended data set is the mean or modal values for the variables in the data set.

Table 18 shows that neural network model two, constructed on the extended data set has drastically different partial effects of the explanatory variables on reenlistment than did neural network model one, which was constructed from a sound theoretical model. Although some changes could and should be expected from adding variables to a model, the size and magnitude of the changes is disconcerting. For example, the effect of being African-American rather than Caucasian for the base case individual, goes from essentially no effect to an increase in the probability of reenlistment of over 44 percent, simply by adding variables to the model. While some change could be expected, this size of change is suspicious.

Another inconsistency in neural network model two is the effect on reenlistment attributed to MILHOUR. This variable was added to the set of explanatory variables merely to add noise to the data set, but the neural network model implies that adjusting the time of day that a member took the survey by one hour later increased his chances of reenlistment by over 19 percent.

TABLE 18: COMPARISON OF THE PARTIAL EFFECTS OF INDIVIDUAL VARIABLES ON THE PROBABILITY OF REENLISTMENT, WITH RESPECT TO THE BASE CASE INDIVIDUAL, FOR NEURAL NETWORK MODELS ONE AND TWO

VARIABLE	BASE CASE	NEURAL NETWORK MODEL ONE	NEURAL NETWORK MODEL TWO
CIVJOB	1	+0.1%	No effect
ENTRYAGE	19.25	No effect	No effect
E4	E3	+9.5%	+43.0%
E5/6	E3	+12.5%	+44.1%
BLACK	WHITEOTH	+0.1%	+44.2%
HISPANIC	WHITEOTH	No effect	+38.1%
SWC	SNC	+0.1%	+37.2%
MNC	SNC	+25.5%	+28.0%
MWC	SNC	+3.5%	+44.3%
TECOCC	0	No effect	No effect
HSCERT	HSDEG	+2.5%	No effect
FACTOR1[2]	0.0097	+0.1%	+35.2%
FACTOR2[3]	0.0091	+0.1%	+41.3%
SPACTIVE	0	***[1]	+31.0%
SEATIME	27	***	+17.5%
OSEATIME	10	***	+16.1%
INCOME	14,000	***	+17.0%
PCSMOVE	2	***	+17.2%
MOMSED	12	***	+16.4%
OFDTYJOB	0	***	+15.2%
CIVJOBOF	1	***	+1.5%
MILHOUR	1200	***	+19.3%
NUMENLST	1	***	+17.1%
DEBT	3	***	+19.0%

Notes: [1]Those variables noted with *** are not included in neural network model number one. [2]Satisfaction with military pay and benefits. [3]Satisfaction with the military lifestyle.

Additionally, several of the added variables have questionable signs. SEATIME, OSEATIME, PCSMOVE, and CIVJOBOF should theoretically all have negative signs; an increase in any of these areas should **decrease** the probability of reenlistment, rather than increase it as neural network model two indicates.

The model developed using the extended data set which includes variables that have no theoretical purpose in the model (neural network model two) presents problems for a policy analyst. If the only problem at hand is prediction then neural network model two is slightly better than the other two models. However, if policy implications are to be determined from the model, neural network model number two, and by extension any model developed without a sound underlying theoretical model, should not be used for policy analysis.

The following chapter concludes this thesis and makes recommendations for follow-on research concerning the use of neural networks in the military manpower and personnel analysis area.

VIII. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

This thesis compared two neural network models and a logistic regression model at the task of predicting reenlistment in the Navy. Reenlistment behavior was modeled for males in the ranks of E-3 to E-6 using 17 variables which were classified into demographic/personal, military characteristics, perceived probability of civilian employment, educational level, and satisfaction with military life and military benefits. Two subsamples were created from the 1985 DoD Officer and Enlisted Personnel Survey; a training sample consisting of 680 observations, and a testing sample consisting of 100 observations.

The neural network models were constructed using NeuralWare software and its default settings, with two hidden neurons in one single hidden layer. Neural network model one was compared to a logistic regression model developed at the Naval PostGraduate School, by George Thomas and Kathryn Kocher. The two models were constructed using the same variables.

At the task of predicting reenlistment the two models created using the same variables performed in a very similar manner. Both models correctly predicted 71 out of the 100 reenlistment decisions in the testing data set. In addition, both models correctly predicted the same number of members who would reenlist, and who would leave the Navy. The logistic regression model had a slightly higher simulation R^2 (.0836) than did the neural network model (.0644), but this did not affect the predictive ability of the neural network model.

For those concerned only with the task of prediction, neural network model one performed as well as did the logistic regression model. However, military manpower and personnel analysts are often more concerned with the policy implications that a model may suggest, rather than simply the predictive power of the model. That is, they are more concerned with what the partial effects of policy variables are, than with how well the model predicts overall.

Neural network model one was found to be deficient as a tool for policy analysts. It ignored those variables which changes in policy can affect, and ascribed most of the effects on reenlistment to those variables in the demographic/personal category which policy changes cannot effect. Neural network model one implies that, for a base case or "typical" individual, improvements in those areas which make up military lifestyle and military benefits, and are likely to lead to higher scores on the composite satisfaction variables, have no effect on the probability of that member's reenlistment.

Another deficiency of both neural network models is the lack of a statistical test for the significance of either individual variables or the model as a whole. This deficiency does not allow the researcher to test hypotheses about the statistical significance of an estimated model or the explanatory variables. For example, when using logistic regression, often there are cases where a change in an explanatory variable will have an effect on the output variable (in this case reenlistment), but the input variable is found not to be statistically significant at some cutoff level. In the neural network models there may be variables which have an estimated effect on reenlistment, yet from a statistical view a null hypothesis of no effect would be supported; there is no way to know this from the results of the neural network model. This is not a serious problem for those researchers concerned with only the predictive capability of a model, but it does present problems for researchers who wish to make policy recommendations based on the model.

Some neural network literature suggests that the "kitchen sink" approach to developing a neural network model is often appropriate [Ref. 7]. That is, if there is no apparent underlying economic model, or if the researcher is unfamiliar with the problem to be modeled, then the neural network model should initially include all the variables in the data set to be examined. The neural network should then be allowed to determine which variables or combinations of variables will affect the output variable. This methodology is in contrast with basic econometric procedures [Ref. 16.] This thesis tested the "kitchen sink" method of model building by adding variables to the original neural network model, some of which had a theoretical background for predicting reenlistment, and some of which were noise for the neural network model to filter. Neural network model two did as well as both the logistic regression model and neural network model one at prediction, but was found to be deficient for policy applications.

B. POLICY IMPLICATIONS

This thesis showed that although neural networks have promise as tools for analysts in the military manpower and personnel field, they cannot yet be used alone for modeling. Neural networks do have applications in these fields, but they should not be used as replacements for more traditional methods of data analysis.

Neural networks have shown promise as predictors. The literature reviewed in Chapter III was nearly unanimous in its support for the use of neural networks as forecasting tools. Although the data set in this thesis yielded a neural network little better at predicting reenlistment than a logistic regression model, the use of neural networks alongside more traditional models as predictors is warranted in other situations not well suited to traditional methods.

The use of neural networks to explain the partial effects of changes in variables should be approached with extreme caution. The lack of statistical tests for evaluating the significance of individual variables or the model as a whole is a major drawback to the use of neural networks. At this time it is recommended that neural networks not be used for developing models to be used for policy analysis.

C. RECOMMENDATIONS

As with most empirical studies, this thesis leaves room for further research. Some recommendations for follow-on research examining the use of neural networks in the manpower and personnel analysis field are discussed below.

One area of research which should be pursued is the comparison of neural network models produced by two different neural network programs. This question is suggested by the widely different results of the neural network and the logistic regression models discussed in this thesis. The policy implications of differing model results from different types of software need to be explored.

Another area of research yet unexplored is whether the results obtained by a researcher using a neural network model can be duplicated by a follow-on researcher. Because the initial starting weights of a neural network are set randomly, is there a way to duplicate the construction of a neural network model so that follow-on researchers can attempt to improve on previous research? The lack of capability to duplicate research would decrease the usefulness of neural networks for the military manpower and personnel analyst.

Further research into the use of neural networks in areas where traditional methods of modeling are weak is also warranted. The problem of modeling reenlistment behavior has been extensively researched, and has been explained quite well using logistic regression. A neural network showed little advantage over a traditional form of data analysis. However, areas exist where traditional methods of modeling are weak. Examples of these weakly modeled domains are those areas such as small data sets, data sets where the dependent variable takes on large numbers of one response and small numbers of another, and data sets where the candidate explanatory variables are all highly correlated to each other. Further research should be done to determine if neural networks may be able to improve modeling in those areas.

Finally, the use of neural networks in areas where it's claimed they are strong should be evaluated. The use of neural networks should be examined in areas where relationships between dependent and independent variables are unknown. In addition, evaluations should be done to determine if researchers with no statistical background can use neural networks effectively as modeling tools. Neural network software makers claim that neural networks are at their strongest in these areas. Neural networks should be applied to data sets with many variables and the resulting models examined to determine if they make sense intuitively.

In summary, neural networks show some promise as tools for the military manpower and personnel analyst. They are a state-of-the-art technology on which millions of dollars of research and development is being spent (much of it at government expense). Neural networks are innovative tools that show some potential for applications in the future. However, researchers should proceed with caution in the use of neural networks, using them alongside more traditional modeling methods for the near future.

LIST OF REFERENCES

1. "Neural Networks Pay Off," Federal Computer Week, 15 February, 1993.

2. NeuralWare Sales Brochure Applying Neural Computing in Business, Industry and Government, 1992.

3. Stanley, J., Introduction to Neural Networks, ed Luedeking, S., 2nd ed., (Sierra Madre, CA: California Scientific Software, 1989).

4. Beale, R. and Jackson, T., Neural Computing, An Introduction, 2nd ed., (New York: Adam Hilger, 1991)

5. Minsky, M. and Papert, S., Perceptrons, (Cambridge MA: MIT Press, 1969).

6. McClelland, J. and Rumelhart, D., Parallel Distributed Processing, (Cambridge MA: MIT Bradford Press, 1986).

7. NeuralWare Applying Neural Computing in Business, Industry and Government class notes, taken December 11-14, 1992, Pittsburgh PA.

8. Dickieson, J., and C. Wilkins, "An exploratory Examination of Artificial Neural Networks as an Alternative to Linear Regression," Independent Research and Independent Exploratory Development Programs: FY91 Annual Report, (San Diego CA: Navy Personnel Research and Development Center, 1991) p. 65-71.

9. Wiggins, V. and Engquist, S., "Neural Networks in Airmen Reenlistment Models," presented at the First Annual Conference on Neural Networks in Military Manpower and Personnel Analysis, on February 2-3 at the Navy Personnel Research and Development Center, San Diego, CA.

10. Surkan , A. and J. Singleton, "Neural Networks for Bond Rating Improved by Multiple Hidden Layers," Proceedings of the 1990 International Joint Conference on Neural Networks (ICJNN), Vol. 2, 157-162, 1990.

11. Odom, M. and R. Sharda, "A Neural Network Model for Bankruptcy Prediction," Proceedings of the 1990 ICJNN, Vol. 2, 163-168, 1990.

12. Afifi, A. and Clark, V., Computer Aided Multivariate Analysis, 2nd ed., (New York: Van Nostrand Reinhold, 1990).

13. Yoon, Y. and G. Swales, "Predicting Stock Price Performance," Proceedings of the 24th Hawaii International Conference on System Sciences, Vol. 4, 156-162, 1991.

14 Hill, T., O'Connor, M., and Remus, W., "Neural Network Models for Time Series Forecasts," A working paper presented at the 24th Hawaii International Conference on System Sciences, 1991.

15. Nileen Hunt, et al., 1985 DoD Survey of Officer and Enlisted Personnel, (Defense Manpower Data Center, 1986).

16. Studenmund, A., Using Econometrics: A Practical Guide, 2nd ed. (New York: HarperCollins Publishers Inc., 1992).

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, VA 22304-6145	2
2.	Library, Code 52 Naval Postgraduate School Monterey, CA 93943-5002	2
3.	LT Bradley S. Russell RT 2 Box 2160 Bandon, OR 97411	2
4.	George Thomas, Code AS/Te Department Of Administrative Sciences Naval Postgraduate School Monterey, CA 93943-5001	4
5.	Daniel R Dolk, Code AS/Dk Department Of Administrative Sciences Naval Postgraduate School Monterey, CA 93943-5001	4
6.	Tim Hill E601J College of Business Administration University of Hawaii 2404 Maile Way Honolulu, HI 96822	1